Time-Domain Fluorescence Diffuse Optical Tomography: Algorithms and Applications

Time-Domain Fluorescence Diffuse Optical Tomography: Algorithms and Applications

a dissertation presented
by
Steven Shuyu Hou
to
Harvard School of Engineering and Applied Sciences

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Applied Physics

Harvard University
Cambridge, Massachusetts
September 2014

Thesis advisors: Professors Brian J. Bacskai, Anand T. N. Kumar Steven Shuyu Hou

Time-Domain Fluorescence Diffuse Optical Tomography: Algorithms and Applications

Abstract

Fluorescence diffuse optical tomography provides non-invasive, in vivo imaging of molecular targets in small animals. While standard fluorescence microscopy is limited to shallow depths and small fields of view, tomographic methods allows recovery of the distribution of fluorescent probes throughout the small animal body. In this thesis, we present novel reconstruction algorithms for the tomographic separation of optical parameters using time-domain (TD) measurements. These technique are validated using simulations and with experimental phantom and mouse imaging studies. We outline the contributions of each chapter of the thesis below.

First, we explore the TD fluorescence tomography reconstruction problem for single and multiple fluorophores with discrete lifetimes. We focus on late arriving photons and compare a direct inversion approach with a two-step, asymptotic approach operating on the same TD data. We show that for lifetime multiplexing, the two methods produce fundamentally different kinds of solutions. The direct inversion is computationally inefficient and results in poor separation but has overall higher resolution while the asymptotic approach provides better separation, relative quantitation of lifetime components and localization but has overall lower resolution. We verify these results with simulation and experimental phantoms.

Second, we introduce novel high resolution lifetime multiplexing algorithms which combine asymptotic methods for separation of fluorophores with the high resolving power of early photon tomography. We show the effectiveness of such methods to achieve high resolution reconstructions of multiple fluorophores in simulations with complex-shaped phantoms, a digital mouse

atlas and also experimentally in fluorescent tube phantoms.

Third, we compare the performance of tomographic spectral and lifetime multiplexing. We show that both of these techniques involve a two-step procedure, consisting of a diffuse propagation step and a basis-function mixing step. However, in these two techniques, the order of the two steps is switched, which leads to a fundamental difference in imaging performance. As an illustration of this difference, we show that the relative concentrations of three colocalized fluorophores in a diffuse medium can accurately be retrieved with lifetime methods but cannot be retrieved with spectral methods.

Fourth, we address the long standing challenge in diffuse optical tomography (DOT) of cross-talk between absorption and scattering. We extend the ideas developed from lifetime multiplexing algorithms by using a constrained optimization approach for separation of absorption and scattering in DOT. Using custom designed phantoms, we demonstrate a novel technique allows better separation of absorption and scattering inclusions compared to existing algorithms for CW and TD diffuse optical tomography.

Finally, we show experimental validation of the lifetime multiplexing algorithms developed in this thesis using three experimental models. First, we show the reconstruction of overlapping complex shapes in a dish phantom. Second, we demonstrate the localization accuracy of lifetime based methods using fluorescent pellets embedded in a sacrificed mouse. Third, we show using planar imaging and tomography, the in vivo recovery of multiple anatomically targeted near-infrared fluorophores.

In summary, we have presented novel reconstruction algorithms and experimental methods that extend the capability of time-domain fluorescence diffuse optical tomography systems. The methods developed in this thesis should also have applicability for general multi-parameter image reconstruction problems.

Contents

Citations to Previously Published Work

Chapter 4 contains portions of text from:

Rice, W. L., Hou, S., & Kumar, A. T. (2013). Resolution below the point spread function for diffuse optical imaging using fluorescence lifetime multiplexing. *Optics letters*, 38(12), 2038–2040

Chapters 4 and 5 contain portions of text from:

Hou, S. S., Rice, W. L., Bacskai, B. J., & Kumar, A. T. (2014b). Tomographic lifetime imaging using combined early-and late-arriving photons. *Optics letters*, 39(5), 1165–1168

Chapter 8 contains portions of text from:

Hou, S., Rice, W. L., Hyun, H., Choi, H. S., Bacskai, B., & Kumar, A. T. (2014a). In vivo tomographic fluorescence lifetime multiplexing of organ targeted fluorophores. In *Biomedical Optics* (pp. BM3A–49).: Optical Society of America

Listing of figures

For my parents

Acknowledgments

My advisors: Brian Bacskai and Anand Kumar

Thesis committee: Vahid Tarokh and Marko Loncar

Brad Hyman and members of the Alzheimer's disease research group

David Boas and members of the Martinos Center optics division

Optical molecular imaging lab: Bill Rice

Two-photon lab: Michal Arbel, Snow Wang, Ksenia Kastanenka, Hong Xie, Scott Raymond, Kishore Kuchibhotla and Naomi Shakerdge

1

Introduction

Molecular imaging allows the visualization of fundamental biological processes such as gene expression and protein interactions in the whole body of small animals (Massoud & Gambhir, 2003; Bremer et al., 2003). Optical methods for molecular imaging use visible to near-infrared light to probe intrinsic optical properties, absorption and scattering or the distribution of fluorophores in tissue. Optical methods can be distinguished from other molecular imaging modalities such as PET, SPECT, MRI and CT as they are relatively cheap, simple to implement and

portable. Secondly, a multitude of molecular probes which are used in microscopy applications can be directly imported for use in whole body imaging. Thirdly, optical methods provide excellent capabilities for multiplexing using either spectral or fluorescence lifetime contrast.

Optical molecular imaging techniques have been widely applied to study disease models in mice. Studies have investigated various cancer related processes such as tumor growth, metastasis formation and gene expression (Weissleder & Ntziachristos, 2003; Montet et al., 2007). Using fluorescent proteins such as GFP has allowed monitoring of surface tumors (Garofalakis et al., 2007; Rice & Kumar, 2014) while the development of near-infrared fluorescent proteins (iRFP) (Shcherbakova & Verkhusha, 2013) allows the potential of imaging deep seated tumors in vivo. Fluorescence tomography has also been applied to the brain for imaging of Amyloid-β plaques in Alzheimer's disease (Hyde et al., 2009) and matrix metalloproteinase activity after stroke (Klohs et al., 2009).

In this thesis, we focus on both the development of reconstruction algorithms for time-domain optical molecular imaging using both fluorescence and intrinsic optical properties as contrast. We also study the application of these techniques in liquid and gel phantoms, phantoms with complex shaped inclusions and living mice. We describe below the various kinds of information that can be obtained from time-domain optical tomography measurements and how this information can be used for image reconstruction.

Existing fluorescence tomography systems are mainly based on continuous wave (CW) excitation. Although time domain methods use more expensive instrumentation and are experimentally more complex, they offer several advantages compared to CW methods. Time domain measurements contain comprehensive information about the biological medium being probed. The temporal point spread function (TPSF) measured at the surface of a turbid medium can be temporally divided into an early portion and a late portion. Early arriving photons have been shown to provide higher resolution than CW measurement and late arriving photons (Chen et al., 2000;

Turner et al., 2005; Niedre et al., 2008). Despite its more diffusive nature, late arriving photons provide information about the intrinsic fluorescence lifetime. Although recovered lifetime values may become distorted due to propagation through a medium (Patterson & Pogue, 1994; Hutchinson et al., 1995; O'Leary et al., 1996), in (Kumar et al., 2005), it was shown that the condition for exact recovery of the intrinsic lifetime from asymptotic TD measurements is based on the diffuse time scale of the medium. In chapter 4, we extend this previous analysis by introducing a new time constant dependent on both the intrinsic lifetime and diffuse time scale which determines the start of the asymptotic regime. We also compare the conditioning of the inverse problem for multiple early and late time points in the case of a single fluorophore. The single fluorophore reconstruction problem can be readily extended to multiplexing multiple fluorophores with discrete lifetimes.

Multiplexing allows the simultaneous tracking of multiple molecular processes and the study of the interactions between different processes. In fluorescence tomography, multiplexing of fluorophores is enabled through separation of the decay portion of a measured TD signal into individual components corresponding to each fluorophore. Since multiplexing of multiple lifetimes has already been successfully applied in fluorescence lifetime imaging microscopy (FLIM), there exist many contrast agents for FLIM which can be translated directly for tomographic imaging (Bastiaens & Squire, 1999; Selvin, 2000; Berezovska et al., 2003). The development of reliable lifetime based reconstruction techniques which could accurately reconstruct the fluorescence yield of individual dyes would allow for measurements which were not previously possible in whole body imaging. These measurements include the detection of molecular interactions using fluorescence resonance energy transfer (FRET) (Jares-Erijman & Jovin, 2003) deep in tissue and the detection of an activatable probe in an activated state over a stronger fluorescence background signal through shifts in lifetime (Goergen et al., 2012).

The tomographic separation of fluorophores using lifetime as contrast has previously been

performed using two approaches. In the direct inversion approach, late time gates at every detector location are combined into a measurement vector and inverted using the single time domain weight matrix which relates fluorescence yield to the measurements (Chen et al., 2011). In the second approach, the time domain weight function is decomposed into spatial and temporal terms. A two step algorithm for reconstruction is then employed which performs a linear multi-exponential fit to obtain the decay amplitudes followed by a separate inversion of CW weight matrices applied on the decay amplitudes (Kumar et al., 2006). In chapter 4, we express the spatio-temporal factorization in matrix form and show that the inverse operators corresponding to the two step approach is mathematically distinct from direct inversion of the TD data. We explore the different forms of the inverse operator for each approach and show the implementation differences between the two forms. We analyze the separation capability of both methods by defining a new quantitative measure for cross-talk between lifetime components. We compare the ability of each approach to recover the relative amounts of two overlapping fluorophores. These simulation results have direct applicability when fluorophores of different lifetimes are in close proximity such as the case with FRET. We further validate our results in an experimental phantom.

While chapter 4 demonstrates that the asymptotic approach results in superior performance for separation of fluorophores, a potential downside of the approach is that its applicability is limited to time points in the asymptotic regime. Such an approach leaves out time points around the peak of the TPSF which have the highest signal-to-noise ratio and time points at the start of the TPSF whose weight matrices are better conditioned than weight matrices for all other time points. In Chapter 5, we develop hybrid methods which combine the resolution of early photons with the complete separation of fluorescence yield offered by the asymptotic approach applied to late arriving photons. Simulations and experimental phantom results are presented comparing the hybrid methods to asymptotic and early photon methods alone.

While we have shown that lifetime can be used to distinguish fluorophores, another method to

characterize a fluorophore is by its excitation and emission spectrum. Spectral contrast has been used extensively in microscopy and has found application in fluorescence tomography, mainly for the separation of autofluorescence from fluorophores of interest (Mansfield et al., 2005). Spectral methods have also been used in diffuse optical tomography for separation of oxy and deoxy-hemoglobin (Boas et al., 2001a). In chapter 6, we compare lifetime and spectral methods for the multiplexing of fluorophores. We make the observation that for both types of imaging, the full weight matrix can be written as a product of a spatially dependent matrix containing CW weight matrices and a matrix containing basis functions (exponential decay functions for lifetime, excitation/emission spectra for spectral). However, lifetime and spectral methods differ in that the order of the matrix product is switched. We explore how this difference in the forward problem affects the separation capability of each method.

In addition to fluorescence imaging, the imaging of optical properties is frequently possible using the same optical tomography system. However, there exist significant differences between the optical property and fluorescence reconstruction problems. While fluorescence tomography in the case of discrete and known lifetimes is a linear inverse problem, optical property tomography is nonlinear inverse problem and image reconstruction requires linearization using a perturbative approach (O'Leary et al., 1995). This step adds to the complexity of the reconstruction algorithm and the validity of the linearized model must be taken into account. An iterative scheme for non-linear reconstruction has also frequently been applied (Schweiger et al., 2005). Like for fluorescence tomography, optical property reconstructions can also benefit significantly from using time domain measurements. Due to non-uniqueness, CW cannot be used to separate absorption and scattering (Arridge, 1999). Reconstruction using time domain data can improve the separation of absorption and scattering although a significant cross-talk component remains between the two parameters (Gao et al., 2002).

In chapter 7, we extend the concepts developed in previous chapters for the separation of ab-

sorption and scattering. We investigate inversion methods which may perform worse in overall error but perform better in separating different parameters using measures such as cross-talk. Since a spatio-temporal factorization can no longer be performed on the time domain weight matrix as in the fluorescence case, we consider a more general approach to improve separation of the parameters of interest. We frame the inverse problem using a Bayesian approach and derive a new estimator which minimizes for mean square error while also placing constraints on inter-parameter cross-talk. We compare this estimator with estimators based on minimum error for optical property reconstruction. We also apply this method for the separation of fluorophores for TD measurement data outside the asymptotic regime.

Even though the development of reconstruction algorithms is important for progress in the field of optical tomography, a parallel challenge remains producing reliable and consistent experimental results in small animal applications. (Raymond et al., 2010) demonstrated an in vivo lifetime based tomography application using dyes with varying specificity for organs. We expand on this previous work by first constructing complex shaped phantoms with known shapes and locations and performing tomographic reconstruction. We also verify the localization accuracy of lifetime based methods in a sacrificed mouse. Finally, we perform lifetime screening and tomography measurements with dyes targeting specific organs.

1.1 Overview

In summary, this dissertation presents novel reconstruction algorithms for time domain measurements of fluorescence and optical property imaging as well as experimental reconstructions in phantoms and in vivo applications. In chapter 2, optical tomography theory is presented including the physical bases for the optical parameters which provide contrast, light propagation models and the solution of ill-conditioned inverse problem. In chapter 3, the instrumentation and exper-

imental procedure are described along with system calibration measurements, novel experimental techniques and new acquisition software. In chapter 4, we analyze the asymptotic and direct inversion approaches for lifetime tomography. In chapter 5, we develop new hybrid reconstruction methods which combine early and late time gates. In chapter 6, we compare lifetime and spectral multiplexing for the separation of multiple fluorophores. In chapter 7, we present a constrained optimization approach for parameter separation in general optical tomography problems. Finally in chapter 8, we validate lifetime based tomography in a complex shaped phantom, sacrificed mouse and in vivo.

2

Theory

Fluorescence diffuse optical tomography can provide a non-invasive method to image the in vivo 3D distribution of molecular probes deep in tissue with applications in preclinical research (Ntziachristos et al., 2005; Leblond et al., 2010). While some general features of an embedded object may be obtained from planar images at the surface of a diffuse medium, tomographic reconstruction methods must be used to determine the spatial location of an object in all three dimensions or extract quantitative information about physical parameters of interest. The reconstruction of

optical parameters from multiply scattered light is a challenging problem and requires accurate models for light propagation in tissue and an understanding of the assumptions that underlie different reconstruction algorithms.

In this chapter, we consider the various optical parameters which provide contrast in optical tomography and their physical bases. We provide an overview of the commonly used models for light propagation in tissue. We review methods of expressing measurement data and unknown optical parameters as linear inverse problems with an emphasis on the problem of fluorescence lifetime multiplexing. We finally review standard techniques to solve ill-conditioned linear inverse problems.

2.1 Optical Parameters in Optical Tomography

An overview of the three main optical parameters, absorption, scattering and fluorescence which we use for imaging is presented in this section.

2.1.1 Absorption

Absorption causes the attenuation of light as it propagates through a medium. Through absorption of a photon, an absorbing molecule or chromophore moves from its electronic ground state to an electronic excited state. The molecule can subsequently return to its ground state through dissipation by heat or emission of a photon through fluorescence or phosphorescence.

The effect of absorption can be modeled statistically by considering a non scattering medium with a single type of absorber. Each absorber can be characterized by its cross sectional area σ which represents its absorbing capacity. For a density of absorbers, N and thickness dx, the relative

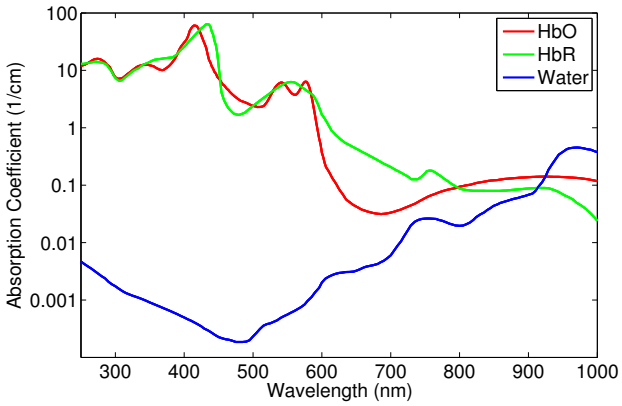

Figure 2.1: The absorption spectra for the three main chromophores in tissue for visible to near-infrared wavelengths (Data taken from (Prahl, 2001)).

reduction in light intensity I is equal to the fractional area occupied by the absorbers:

$$dI/I = -\sigma N dx \tag{2.1}$$

Integrating for thickness 0 to x leads to the Beer-Lambert Law:

$$I(x) = I_0 exp(-\mu_a x) \tag{2.2}$$

where $\mu_a = \sigma N$ in units of cm^{-1} is called the absorption coefficient and can be interpreted as the probability of photon absorption per unit infinitesimal length. μ_a can equivalently be expressed as: $\mu_a = \epsilon C$ where ϵ is the extinction coefficient in units of M^{-1} cm^{-1} and C is the concentration in units of M.

For tissue, the main chromophores for near-infrared wavelengths are oxygenated (HbO) and deoxygenated (HbR) hemoglobin, and water. Other minor contributors include melanin, fat and bilirubin. The absorption spectra of the three main chromophores are shown in Fig. 2.1 for visible to near-infrared wavelengths.

From Fig. 2.1, the 'optical window' corresponding to wavelengths from 650 nm to 900 nm is the region of the spectrum where the chromophores have low absorption. Non-invasive optical tomography is typically performed in this spectral range to maximize tissue penetration. For shorter wavelengths, hemoglobin absorption dominates while for longer wavelengths water absorption dominates resulting in much reduced light transmission in both cases. Measurements of temporal dynamics of HbO and HbR in this wavelength range can provide important physiological information about blood oxygenation and volume and have been applied for stroke imaging (Culver et al., 2003) in small animals and breast cancer imaging (Fang & Boas, 2009) and functional brain imaging (Boas et al., 2001b) in humans. Due to the unique absorption spectra for both chromophores, estimates for their concentrations can be obtained from absorption measurements at two wavelengths.

2.1.2 Scattering

Scattering causes the broadening of the light distribution in tissue and manifests itself as changes in direction of the photon through collisions inside a medium. In the case of a highly scattering medium such as biological tissue, photons will quickly undergo enough collisions that their directions become random. In our propagation models, photon energy is modeled as being preserved after collisions (elastic scattering). Similar to absorption, a measure for scattering can be defined in a statistical manner. The scattering coefficient μ_s in units of cm^{-1} is the probability of a photon being scattered per unit infinitesimal length.

Scattering arises from refraction index mismatches within a medium and depends on many factors including the sizes of the scattering particles, the wavelength of light and on the indices of refraction of the particles and the medium (Niemz, 2007). Scattering due to particles of any size can be approximately modeled using Mie theory which describes the scattering of light by spheres. For particles of size comparable to the light wavelength, scattering as predicted by Mie theory

shows a wavelength dependence of $\sim \lambda^{-x}$ with $0.4 \leq x \leq 0.5$. For scattering particles much smaller than the wavelength of light, Mie theory can be reduced to Rayleigh theory in which case the scattering shows a wavelength dependence of $\sim \lambda^{-4}$. It has been found that the wavelength dependence measured in tissue cannot be explained by either Mie theory or Rayleigh theory alone (Niemz, 2007). Alternatively, tissue scattering can be modeled considering separate contributions of Mie and Rayleigh scattering (Jacques, 2013).

A scattering phase function, $P(\hat{s}, \hat{s}')$ is used to characterize the probability distribution of a photon's initial direction $\hat{s}$ and direction after scattering $\hat{s}'$. Tissue is typically assumed to be an isotropic medium so that the phase function is independent of the absolute $\hat{s}$ and $\hat{s}'$ of the photon and only depends on the angle between them $P(\hat{s}, \hat{s}') = P(\hat{s} \cdot \hat{s}')$. A commonly used phase function is the Henyey-Greenstein function:

$$p(cos(\theta)) = \frac{1 - g^2}{2(1 + g^2 - 2gcos(\theta)^{3/2})} \tag{2.3}$$

A measure for the direction of scattering in a medium is the anisotropy, g which can be found by taking the average of the cosine of the scattering angle. Scattering in tissue has been shown to be highly forward scattering with $g > 0.9$ (Cheong et al., 1990).

For wavelengths in the optical window, the cell constituents which are responsible for scattering are mainly intracellular components such as mitochondria and nuclei (Mourant et al., 1998).

2.1.3 Fluorescence

Fluorescence is the emission of light from a molecule following absorption and can be described by the following sequence of steps (Lakowicz, 2006). First, a fluorescent molecule or fluorophore absorbs a photon and goes from the electronic ground state, S_0 to an electronic excited state (usually S_1 or S_2). Through collisions with surrounding molecules, the fluorophore loses energy non-

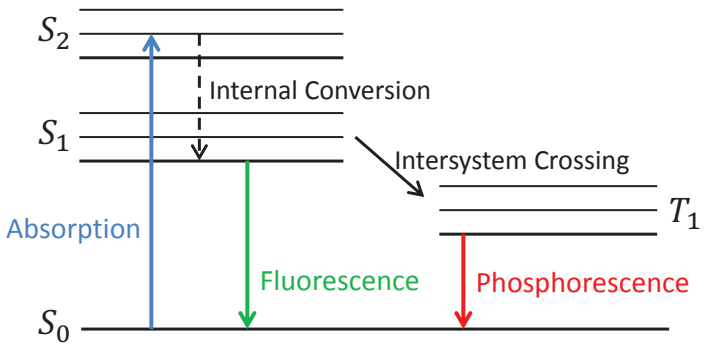

Figure 2.2: A Jablonski diagram showing the various transitions which occur after excitation of a fluorophore.

radiatively and goes to the lowest vibrational energy level of S_1 in a process called internal conversion. From this state, the fluorophore can return to some vibrational energy level of S_0 either through a non-radiative process characterized by rate k_{nr} or through emission of a photon characterized by the fluorescence decay rate k_f. Because the fluorophore has lost some of its vibrational energy after excitation and returns to a higher energy vibrational state than it started in, the emitted photon is red shifted relative to the absorbed photon. This phenomenon is known as the Stokes shift. The entire process can be illustrated schematically with the Jablonski diagram shown in Fig. 2.2.

Two important parameters that characterize a fluorophore are its lifetime and its quantum yield. Quantum yield is a measure for the brightness of a fluorophore and is defined as the ratio of the number of emitted photons to the number of absorbed photons. It can be expressed in terms of the rate constants, k_{nr} and k_f:

$$Q = k_f / \left(k_{nr} + k_f\right) \tag{2.4}$$

Fluorescence lifetime is an important parameter only accessible by time resolved measurements. Following a short pulsed excitation of a sample, an initial population of n_0 excited states is cre-

ated. The temporal dynamics of the number of excited states n can be described by the following equation:

$$dn(t)/dt = -(k_{nr} + k_f)n \tag{2.5}$$

By solving (2.5), the number of molecules in the excited state can be expressed as:

$$n(t) = n_0 exp(-(k_{nr} + k_f)t) \tag{2.6}$$

Since the measured fluorescence intensity $I(t)$ is directly proportional to $n(t)$ and given the fluorescence lifetime is defined as the inverse of the rates of decay $\tau = 1/(k_{nr} + k_f)$, $I(t)$ can also be written as a exponential decay function:

$$I(t) = I_0 exp(-t/\tau) \tag{2.7}$$

The lifetime represents the average amount of time a fluorophore with a single exponential decay function stays in its excited state before a photon is emitted and the fluorophore returns to its ground state:

$$< t >= \frac{\int_0^\infty tI(t)dt}{\int_0^\infty I(t)dt} = \tau \tag{2.8}$$

A major advantage of using fluorescence lifetime as a quantitative measure is that it is generally independent of concentration. Due to the relatively long duration of typical fluorescence lifetimes (ns) compared to absorption which occurs on the fs time scale, lifetime can also be sensitive to many biochemical processes not available from absorption measurements such as environmental changes in viscosity, pH, polarity and solvent relaxation, molecular interactions, energy transfer and quenching (Lakowicz, 2006).

2.2 Light Propagation Forward Model

In the forward problem, the light intensity at detector positions on the surface of a turbid medium is determined from a given distribution of the absorption coefficient, scattering coefficient and fluorescence yield throughout the medium. The function relating the optical parameters to the measurements can be expressed in terms of fundamental solutions to photon transport equations.

2.2.1 Radiative Transport Equation

A general model for propagation of photons in biological tissue is the radiative transport equation (RTE). The RTE can be derived based on energy conservation within each volume element of phase space while neglecting effects such as polarization and coherence. The quantity of interest is the radiance L measured in units of W m^{-2} sr^{-1} and is a function of six independent variables $(r, \hat{s}, t)$. The time-dependent RTE, expressed in integro-differential form is given by (Arridge, 1999):

$$\left[\frac{1}{v}\frac{\partial}{\partial t} + \hat{s}\cdot\nabla + \mu_t\right] L(r,\hat{s},t) = \mu_s \int_{4\pi} L(r,\hat{s}',t)P(\hat{s}\cdot\hat{s}')d\hat{s}' + S(r,\hat{s},t) \tag{2.9}$$

where v is the speed of light in the medium, μ_t is the transport coefficient equal to the sum of μ_s and μ_a, $S(r,\hat{s},t)$ is the photon source, $L(r,\hat{s},t)$ is the fluence, $\hat{s}$ is the unit direction vector and $P(\hat{s}\cdot\hat{s}')$ is the scattering phase function, frequently modeled by the Heyney-Greenstein function (2.3).

The RTE cannot be solved analytically for any practical geometries. A common numerical method of solving the RTE is with Monte Carlo simulation (Wang et al., 1995). In this dissertation, the RTE is primarily solved using the Monte Carlo approach. A brief overview about Monte Carlo methods is given in section 2.2.2. Other numerical methods for solving the RTE use higher order approximations such as expansion in spherical harmonics (P_N) (Aydin et al., 2002) and

discrete ordinates (S_N) (Abdoulaev & Hielscher, 2003). The first order spherical harmonics approximation (P_1) results in widely used diffusion equation (section 2.2.3).

2.2.2 Monte Carlo Simulation

Monte Carlo software for photon propagation has seen significant progress in recent years including improvements in computation time from GPU based implementations (Prahl et al., 1989; Wang et al., 1995; Fang & Boas, 2009). In the Monte Carlo approach, a random walk for N photons is performed. Since the signal to noise ratio of a Monte Carlo simulation depends on $\sim \sqrt{N}$, a large number of photons must be simulated to produce reliable estimates. The standard rules for incrementally moving a photon through a medium are outlined by (Wang et al., 1995) and are used in the Monte Carlo simulations (Boas et al., 2002) presented in later chapters . We briefly describe them below.

Absorption is modeled by assigning a weight to every simulated photon. As a photon propagates through the medium its weight is reduced by a factor of $exp(-\mu_a L)$ where L is the distance traveled. The distance a photon travels between scattering events is determined by sampling an exponential distribution based on the scattering coefficient at the location of the photon. The new direction of the photon is determined from sampling the scattering phase function. Fresnel's equations are used to find the probability of internal reflection at medium boundaries. The random walk for a photon ends either when it leaves the medium or it makes contact with a set of predefined detectors.

2.2.3 Diffusion Equation

The diffusion equation provides a computationally tractable model for light propagation but is not generally applicable like the RTE. The time-dependent diffusion equation is given as (Wang &

Wu, 2007):

$$\frac{\partial U(r,t)}{v\partial t} + \mu_a U(r,t) - \nabla \cdot [D\nabla U(r,t)] = S(r,t) \tag{2.10}$$

where $U(r,t)$ is the fluence measured in units of W cm^{-2} which is the integral of the radiance $U(r,t) = \int_\Omega L(r,\hat{s},t)d\hat{s}$, $\mu_s' = (1-g)\mu_s$ is the reduced scattering coefficient, where $1/\mu_s'$ is the distance a photon travels before scattering has made its direction isotropic and $D = 1/(3(\mu_a+\mu_s'))$ is called the diffusion coefficient.

Unlike the RTE, the diffusion equation has analytical solutions for certain homogeneous geometries. For example, in an infinite homogeneous medium for source, $S(r,t) = \delta(r,t)$, the solution is in the form of a damped exponential with an asymptotic decay rate $\mu_a c$:

$$\Phi(r,t) = \frac{c}{(4\pi Dct)^{3/2}} exp\left(-\frac{r^2}{4Dct} - \mu_a ct\right) \tag{2.11}$$

For more complex geometries, a finite element method can be used to numerically solve the diffusion equation (Arridge et al., 1993).

The diffusion approximation is only valid when the radiance has become nearly isotropic due to multiple scattering events. The general conditions for validity of the diffusion equation are detailed in (Boas, 1996) and summarized below.

First, the transport albedo $c = \mu_s'/(\mu_s' + \mu_a)$ should be close to unity or $\mu_s' \gg \mu_a$. This is generally satisfied in tissue where μ_s' can be as much as 100 times greater than μ_a . However, in small animal imaging there exist highly vascularized organs such as the heart or the liver where literature values suggest μ_a can be larger than 1 cm^{-1} (Krainov et al., 2013; Alexandrakis et al., 2005). For these regions of high albedo, it is unclear if the diffusion approximation holds with sufficient accuracy.

Second, the scattering should not be too anisotropic. From (Cheong et al., 1990) where tissue g has been measured to be between 0.68 and 0.99, this condition is generally satisfied.

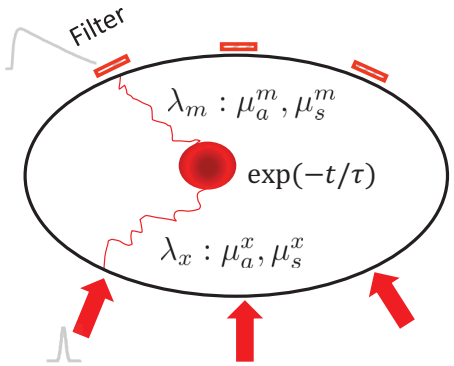

Figure 2.3: Schematic showing the three components of fluorescent light transport in a scattering medium: propagation of the excitation light at λ_x, response of the fluorophore to excitation ($exp(-t/\tau)$) and propagation of the fluorescence light at λ_m. Also shown is the broadening of the temporal response due to the medium compared to the excitation pulse.

Finally, sources should be at least one random walk length ($1/\mu_s'$) from detectors. For a standard $\mu_s' = 10\,\mathrm{cm}^{-1}$, the random walk length is $1/\mu_s' = 1$ mm. This condition should hold easily for small animal imaging in the transmission geometry.

2.2.4 Fluorescence Light Transport

Photon propagation in a medium with fluorophores can be modeled as propagation of the excitation light from the source to fluorophores inside the medium and propagation of the emitted fluorescence from the fluorophore throughout the medium (shown schematically in Fig. 2.3). This process can be described by coupled RTE's and a rate equation relating the excitation field to the temporal response of the fluorophore (Klose, 2009; Hutchinson et al., 1995, 1996).

First, the excitation field can be modeled as before using a RTE for photons at the excitation wavelength λ_x:

$$\left[\frac{1}{v}\frac{\partial}{\partial t} + \hat{s}\cdot\nabla + (\mu_t^x + \mu_{af}^x)\right] L(r,\hat{s},t) = \mu_s^x \int_{4\pi} L(r,\hat{s}',t) P(\hat{s}\cdot\hat{s}')d\hat{s}' + S(r,\hat{s},t) \tag{2.12}$$

where the absorption due to the fluorophore, μ_{af}^x is separated from the background absorption of

the medium. The fluorophore response due to the excitation field can be obtained using the rate equation for excited state fluorophores where the source is the excitation light being absorbed by the fluorophore:

$$dn(t)/dt = -(1/\tau)n + \mu_{af}^{x}\phi(r,t) \tag{2.13}$$

The solution to this equation is given by:

$$n(t) = \mu_{af}^{x}\int_{0}^{t}\phi(r,t')exp\left(-\frac{t-t'}{\tau}\right)dt' \tag{2.14}$$

The fluorophore response multiplied by the quantum yield term Q acts as the source term for the RTE equation at the emission wavelength λ_m:

$$\left[\frac{1}{v}\frac{\partial}{\partial t} + \hat{s}\cdot\nabla + (\mu_t^x + \mu_{af}^x)\right]L(r,\hat{s},t) = \mu_s^x\int_{4\pi}L(r,\hat{s}',t)P(\hat{s}\cdot\hat{s}')d\hat{s}' + Q\mu_{af}^x\int_0^t\phi(r,t')exp\left(-\frac{t-t'}{\tau}\right)dt' \tag{2.15}$$

In the next section, we solve the forward problems for optical properties and fluorescence using Green's function methods.

2.3 Solving the Forward Problem

A general method to solve the RTE or diffusion equation for arbitrary sources and boundary conditions is the application of Green's functions. If the volume Green's function, $G(r,t;r',t')$ is defined as the response to a source $S(r,t) = \delta(r,t)$ where $\delta(r,t)$ is the Dirac delta function, then the solution to an arbitrary source $S(r',t')$ is given by (Case & Zweifel, 1967):

$$U(r,t) = \int_0^t\int_0^\infty G(r,t;r',t')S(r',t')dr'dt' \tag{2.16}$$

We also introduce a symmetry property which is valid for both the RTE and diffusion equation Green's functions. Frequently due to implementation considerations, we seek solutions to an adjoint problem. A standard procedure to find the adjoint solution is by solving the direct problem and applying a reciprocity relation (Case & Zweifel, 1967):

$$G(r, t; r', t') = G(r', -t'; ; r, -t). \tag{2.17}$$

Application of (2.17) can result in significant reduction in the total number of Monte Carlo simulations needed to solve the forward problem. Next, the Green's function methods are used to find linear functions which relate the fluence at detectors to changes in optical properties and fluorophore concentration.

2.3.1 Optical Property

Since optical properties ($\mu_a(r)$, $\mu_s(r)$) are related to the surface measurements through a nonlinear function, the problem is first simplified by linearization using a perturbation approach (O'Leary et al., 1995). This linearization has most commonly been performed for the diffusion equation and the resulting Jacobians for absorption and scattering are presented below. Jacobians based on the RTE are presented in Chapter 6.

Using the perturbation approach, the total parameters, $\mu_a(r)$ and $\mu_s(r)$ are equal to the sum of background terms and perturbation terms:

$$\mu_a(r) = \mu_{a0}(r) + \delta\mu_a(r) \tag{2.18}$$

$$\mu_s(r) = \mu_{s0}(r) + \delta\mu_s(r) \tag{2.19}$$

By substituting (2.18) and (2.19) into the time-dependent diffusion equation (2.10), assuming all first order terms go to zero (Born approximation) and application of (2.16) and (2.17), the perturbation in measured signal ($\delta U_a(r,t)$ and $\delta U_s(r,t)$) due to an absorption and a scattering perturbation can be expressed as:

$$\delta U_a(r,t) = -\int_V G(r_s,r,t) \bigotimes G(r,r_d,t)\delta\mu_a(r)d^3r \tag{2.20}$$

$$\delta U_s(r,t) = -\int_V \nabla G(r_s,r,t) \bigotimes \nabla G(r,r_d,t)\delta\mu_s(r)d^3r \tag{2.21}$$

where r_s and r_d are the source and detector locations respectively and $G(r_s,r,t)$ and $G(r,r_d,t)$ are the source and detector Green's functions. It should be noted that the above equations assume the validity of the diffusion equation and may become less accurate for early time points where photons have undergone minimal scattering events.

2.3.2 Fluorescence

Applying the definition of the Green's function to (2.15) and assuming the emission of photons is due only to absorption at the excitation wavelength while neglecting re-emission due to absorption of photons at the emission wavelength, the fluorescence intensity can be expressed as double time convolutions:

$$U(r_s,r_d,t) = \int_V d^3r \int_0^t dt' \int_0^{t'} dt'' G_{ex}(r,r_s,t'')exp\left(-\frac{t'-t''}{\tau(r)}\right) G_{em}(r_d,r,t-t')\eta(r) \tag{2.22}$$

In this equation, the interactions between emission wavelength photons with the background medium and fluorophores are included in the emission Green's function G_{em}. Here, a new term called fluorescence yield is defined as $\eta = Q\mu_{ax}^f$. Using a more general method for fluorescence

tomography, (Eppstein et al., 2003) has investigated the case when emitted fluorescence can re-excite the fluorophore using a full perturbation approach.

For fluorescence tomography problems investigated in this dissertation, fluorophores are assumed to have known discrete lifetimes τ_n with the unknown parameters being the fluorescence yields associated with each fluorophore η_n. We have found this assumption is generally satisfied for near-infrared dyes and fluorescent proteins in vivo. This allows the fluorescence inverse problem to be solved directly by linear inverse methods for the unknown fluorescence yields.

2.3.3 Asymptotic Fluorescence Lifetime Analysis

Understanding the effects of a turbid medium on a fluorophore's intrinsic lifetime is important for the implementation of lifetime based reconstruction algorithms. This problem was studied in (Kumar et al., 2005, 2006) using asymptotic lifetime analysis and the following discussion is taken largely from these sources. First it was shown that in a scattering medium, the fluorescence intensity can be written as a sum of a diffusive decay term and a fluorescence decay term:

$$U(r_s, r_d, t) = \sum_{i=1}^{n} a_D(r_s, r_d, t) exp(-v\mu_a t) + a_{F_n}(r_s, r_d) exp\left(-t/\tau_n\right) \tag{2.23}$$

where the diffusive term decays with a time constant, τ_D. From (2.23) it can be observed that if fluorescence term decays at a slower rate than the diffusive term ($\tau_n > \tau_D$), then the fluorescence intensity at times after the diffusive decay term has completed its evolution can be used to directly measure the intrinsic fluorescence lifetime in a medium by fitting to an exponential basis function. This condition is verified with simulations in chapter 4.

To apply this concept to the fluorescence tomography reconstruction problem, we consider the time dependent weight term which can be written as a time convolution between an exponential

basis function and a time-dependent background weight function:

$$W(r_s, r_d, r, t) = \int_0^t dt' exp\left(-\frac{(t-t')}{\tau(r)}\right) W_B(r_s, r_d, r, t') \tag{2.24}$$

$$W_B(r_s, r_d, r, t) = \int_0^t dt' G_{ex}(r, r_s, t-t') G_{em}(r_d, r, t') \tag{2.25}$$

For lifetimes longer than the absorption timescale ($\tau_n > \tau_a = 1/(v\mu_a)$), (2.24) can be written in a new form. First, it can be shown that when the Green's function for the RTE is written as a product of a background Green's function G_0 and an exponential decay term, $G(r,t) = G_0(r,t)exp\,(-v\mu_a(t))$, G_0 is only dependent on $\nabla\mu_a$ and is invariant to constant shifts in μ_a. Based on this property, a new reduced absorption Green's function, G_n can be introduced which is the RTE Green's function evaluated at background absorption, $\mu_a - 1/(v\tau)$:

$$G_n(r,t) = G(r,t)\mid_{\mu_a - 1/(v\tau_n)} = G(r,t)\mid_{\mu_a} exp(-t/\tau_n) \tag{2.26}$$

A fluorophore dependent background weight function W_B^n can then be written in terms of reduced absorption Green's functions. This allows the fluorescence measurement from (2.22) to be expressed as the product of a time-dependent amplitude function and an exponential decay term:

$$U(r_s, r_d, t) = \sum_n A_n(r_s, r_d, t) exp\,(-t/\tau) \tag{2.27}$$

$$A_n(r_s, r_d, t) = \int_V d^3r \left[\int_0^t dt' W_B(r_s, r_d, r, t)\right] \eta_n(r) \tag{2.28}$$

For times $t \gg \tau_D$, $A_n(r_s, r_d, t)$ approaches a constant in time and the fluorescence intensity can be

written as a sum of weighted exponential decay functions:

$$U(r_s, r_d, t) = \sum_n exp\left(-t/\tau_n\right) \int_V dr W_B^n(r_s, r_d, r)\eta_n(r) \tag{2.29}$$

where W_B^n are generated from CW reduced absorption Green's functions. It can be seen that in the asymptotic regime the fluorescence intensity at the surface of a scattering medium reduces to the same form as the fluorescent response from direct excitation of a fluorophore.

2.4 Optical Tomography Inverse Problem

The goal of the inverse problem is to reconstruct the spatial distribution of the unknown optical parameters from surface measurements. After discretization of the forward operators, the reconstruction for both optical properties and fluorescence yield are reduced to solving a linear matrix equation:

$$y = Wx \tag{2.30}$$

where y is a known vector containing measurements at combinations of source/detector pairs, time gates and wavelengths and x is an unknown vector containing the optical parameters at every voxel in the medium.

A standard method to solve (2.30) is by the method of least squares where the x which minimizes the residual between the measurement y and Wx is found:

$$x = \arg\min_x \|Ax - y\|^2 \tag{2.31}$$

The solution of (2.31) can be obtained using the generalized inverse or Moore-Penrose pseudoinverse ($A^\dagger$):

$$x_{LS} = A^\dagger y \tag{2.32}$$

The generalized solution in (2.32) has the property that when A is invertible, $A^{\dagger}$ is equal to the inverse matrix, A^{-1}. In the case when multiple solution produce the same minimum residual (for example, in underdetermined systems when an infinite number of solutions, x result in zero residual), the generalized solution corresponds to the minimum norm ($\|x\|$) solution.

Even though the generalized solution in (2.32) is the x that after linear transformation with W best matches the measurement data, in the presence of noise it will not necessarily be a solution with the lowest reconstruction error ($\|x_{recon} - x_{true}\|^2$). In fact, generalized solutions are seldom useful in optical tomographic reconstruction. Due to their high sensitivity to noise, generalized solutions typically show non physical characteristics such as large oscillations. The noise sensitivity of a solution can be analyzed by considering the conditioning of the weight matrix A which describes the amount a change in y produces a corresponding change in the inverted solutions x. The degree of conditioning can be measured using the condition number α which gives a bound on how much the error can propagate from the data to the solution:

$$\frac{\Delta x}{x} \leq \alpha \frac{\Delta y}{y} \tag{2.33}$$

One of the major difficulties associated with all types of diffuse optical imaging is that W has a large condition number. To address this problem, regularization methods are typically used. The process of regularization involves finding solutions to an approximate version of the original inverse problem by incorporation of prior information about the unknowns. A commonly used regularization technique is Tikhonov regularization where a penalty term based on the norm of the solution is added to the least square functional (Bertero & Boccacci, 1985):

$$\Phi(x) = \|Ax - y\|^2 + \lambda \|x\|^2 \tag{2.34}$$

where λ is a real positive number called the regularization parameter. It can be shown that there exists a λ which minimizes the reconstruction error. This property of the algorithm is referred to as semi-convergence.

For a given λ, the solution x that minimizes (2.34) can be written in two equivalent forms:

$$x_{TR} = A^T(AA^T + \lambda I)^{-1}y \tag{2.35}$$

$$x_{TR} = (A^TA + \lambda I)^{-1}A^Ty \tag{2.36}$$

The first form is used for undetermined problems while the second is used for overdetermined problems. Selecting the proper form will reduce the computation time associated with performing the matrix inversion.

One of the main difficulties with ill-conditioned inverse problems is that the λ which produces the solution with the least reconstruction error cannot be determined without knowledge of the true object although there exist certain heuristic methods such as L-curve analysis which can be used to select the value (Hansen, 1992).

To better understand the sources of error when using a regularized inverse, the two terms that result from multiplication of the Tikhonov inverse operator, A_{inv} from (2.35) and (2.36) with a measurement vector containing noise ($y = Ax + n$) is given below (Bertero & Boccacci, 1985):

$$x_{TR} = A_{inv}Ax + A_{inv}n \tag{2.37}$$

From the first term of (2.37), the matrix $A_{inv}A$ is called the resolution matrix of the inverse operator. Each element ij of the matrix represents the response at voxel j due to an object at voxel i in x. Therefore, columns of the matrix contain the point spread function (PSF) for different voxels

in the medium. The difference between the $A_{inv}Ax$ and x is called the approximation error and it quantifies the effect of using a regularized inverse compared to $A^{\dagger}$ when noise is not considered. The second term of (2.37) represents the effect of noise on the inverse operator and is called the noise amplification error. As the regularization parameter is decreased, the resolution matrix approaches the identity matrix and the solution becomes less biased while the error due to the noise term increases. As the regularization parameter is increased, the solution becomes more biased and the corresponding point spread functions in the resolution matrix expand while the noise amplification error is decreased. Selecting a regularization parameter involves finding a balance between these two sources of error.

A well known method to improve computational efficiency while calculating a regularized inverse is to make use of the connection between the singular value decomposition (SVD) and Tikhonov regularization. Given the SVD of A ($A = USV^T$), the solutions in (2.35) and (2.36) can be expressed as (Vogel, 2002):

$$x_{LS} = \sum_{i=1}^{n} \frac{(u_i^T d)}{s_i} v_i \tag{2.38}$$

$$x_{TR} = \sum_{i=1}^{n} \frac{s_i (u_i^T d)}{s_i^2 + \lambda} v_i \tag{2.39}$$

where u_i is the column i of U and v_i is column i of V and s_i are the singular values in S. From (2.38), it can be seen that noise in the measurement data is amplified by small singular values which causes the instability in the least squares solution. From (2.39), Tikhonov Regularization can be viewed as a method to stabilize the solution by reducing the effect of small singular values while retaining the effect of large singular values. Since the singular vectors v_i contain lower spatial frequencies for larger singular values, the regularized solution is a smoothed version of the true object.

The computational advantage of using the SVD becomes apparent when evaluating the inverse

for multiple regularization parameters. In this case, an initial singular value decomposition followed by subsequent application of (2.39) is frequently more computationally efficient than the repeated calculation of the matrix inverse in (2.35) or (2.36).

An alternate approach for solving linear inverse problems is by applying a Bayesian approach. In this case prior information about the object is explicitly modeled with a probability distribution. These methods will be summarized in Chapter 6.

3

Instrumentation and Experimental Procedure

3.1 Introduction

Most early optical tomography systems used optical fibers placed in contact with the sample to deliver light and detect light that has propagated in the sample (Schulz et al., 2006). Fiber-based

systems use a matching medium or compression to conform to a constant sample geometry. Recently, there has been interest in the development of noncontact systems which can avoid experimental errors introduced by fiber/tissue contact and the use of matching mediums (Schulz et al., 2003; Kumar et al., 2008b; Graves et al., 2003). In noncontact systems, a collimated free-space laser beam is delivered to the sample and light is detected with a CCD camera and lens. Despite its flexibility in imaging samples of various sizes, a noncontact system requires accurate methods to capture the 3D surface of the sample.

The methods for illuminating and recording from tissue can be generally divided into either planar and tomographic. Planar imaging is performed in the reflection geometry where wide field sources and detectors are positioned on the same side of the sample. Planar imaging is used to generate a single 2D image with no additional processing. Therefore, there is no explicit information about depth of the object being imaged although objects closer to the surface are more heavily weighted than deeper ones. Tomographic imaging is typically performed in the transmission geometry where point sources and a wide field detector are on opposite sides of the sample. Measurement data for multiple source and detector locations are used in tomographic reconstruction algorithms to recover the 3D distributions of optical parameters. These reconstruction algorithms are based on light transport models discussed in chapter 2.

Optical tomography systems can also be categorized based on their method of excitation: continuous wave (CW) systems use time-independent sources (Ntziachristos & Weissleder, 2001; Zacharakis et al., 2005; Klose & Hielscher, 2003), frequency domain (FD) systems use intensity modulated sources with frequencies in the MHz to GHz range (Milstein et al., 2003; Godavarty et al., 2005) and time-domain (TD) systems use nanosecond to picosecond pulsed sources (Das et al., 1997; Bloch et al., 2005; Kumar et al., 2008b).

CW systems are the simplest to implement and are the least expensive as compared to FD and TD systems. Typical sources for CW imaging are LED's and laser diode's (LD) and detection

can be fiber based or wide-field with a CCD camera. Despite its simplicity, CW is limited by its inability to differentiate absorption and scattering in optical property measurements and fluorescence lifetime and yield in fluorescence tomography. In FD systems, the intensity modulated source creates diffusive waves of the same frequency as the source which propagate in highly scattering media (O'Leary, 1996). The phase shift and amplitude change of the intensity modulated signal at the detectors can be used to determine absorption and scattering perturbations within the medium and recover fluorescence lifetime and yield. Since the response to a laser pulse implicitly contains all modulation frequencies including the CW (zero frequency) component, TD measurements offer the most comprehensive information about the imaging medium. However, due to the popularity of FD systems, TD data is commonly processed through Fourier transform into a few discrete frequencies (Stott & Boas, 2002; Nothdurft et al., 2009). Two aspects of TD data can be highlighted. Early time points represent photons which have undergone fewer scattering events than CW data. Therefore, using early time point data for reconstruction results in better conditioned inverse problems and subsequently higher spatial resolution reconstructions (Leblond et al., 2009). Also, information about intrinsic fluorescence lifetime within a medium are encoded into the late time points of the temporal response (O'Leary et al., 1996).

3.2 System

A time-domain fluorescence diffuse optical tomography system for small animal imaging has been developed in the Optical Molecular Imaging Lab at Massachusetts General Hospital. The device uses a noncontact method for temporal measurements and free-space laser delivery at the surface of the animal. A schematic of our system is shown in Fig 3.1.

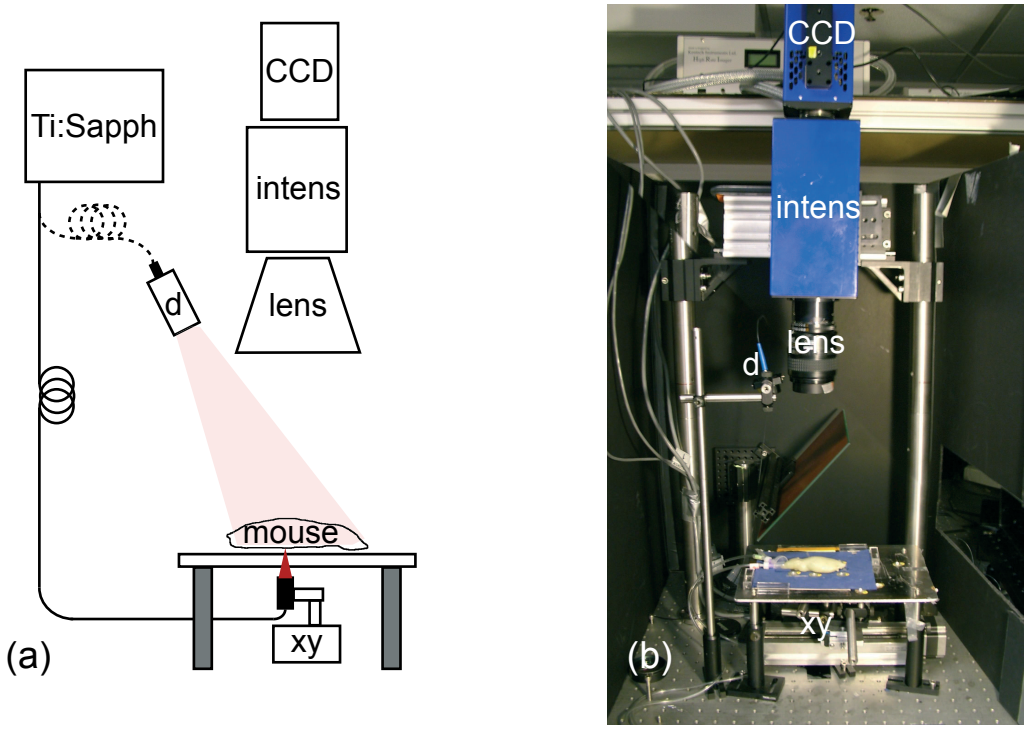

Figure 3.1: Time-domain fluorescence diffuse optical tomography system for small animal imaging. (a) A schematic of the system is shown. For transmission measurements, the source fiber is attached to the translation stage (xy) and scanned across the bottom surface of the mouse. For reflection measurements, the source fiber is attached to an engineered diffuser (d) to provide wide field illumination of the mouse. An intensified CCD camera is used for time-resolved detection. (b) A photograph of the system is shown. To eliminate stray light, a black enclosure is used to surround the system during image acquisition.

3.2.1 Source

For excitation in the near-infrared spectrum, a Ti:Sapphire Laser (Mai-Tai, Newport-Spectra Physics, Moutain View CA) was used with a 150 ps pulse width, 80 MHz repetition rate and tuning range between 700-900 nm. To achieve a pulsed source in the visible spectrum, a photonic crystal fiber (Femtowhite 800, Crystal Fibre, Denmark) was pumped at 800 nm with the Ti-Sapph. Excitation filters that span the range 480-700 nm were used to select out narrow ranges of wavelengths. To control for the amount of input laser power, a rotating polarizer was placed at the output of the Ti-Sapph. Typical output powers were in the range of $\approx$ 10-100 mW. The laser was launched into a 400 µm multi-mode fiber through a fiber collimation package. The output of the fiber was also attached to a collimation package. For reflection imaging, the fiber was moved above the sample and an engineered diffuser was placed in front of the fiber output to provide wide field illumination for an area $\approx$ 5 cm $\times$ 5 cm. For transmission measurements, the fiber output was moved below the imaging platform and attached to a linear translation stage (Velmex Inc., Bloomfield, NY) which allowed for scanning of different x-y positions with a precision of $\sim$ 1 µm. The laser beam was focused to a $<$ 1 mm diameter spot at the imaging plane.

3.2.2 Detection

We first review the operation of an intensified CCD camera (ICCD) for time resolved detection (Selb et al., 2006; Patwardhan & Culver, 2008; Kumar et al., 2008b). ICCD cameras consist of two main components: an image intensifier which acts as a fast optical shutter and a CCD camera. An image intensifier can be described in terms of three elements: a gateable photocathode, a microchannel plate (MCP) for electron multiplication and a phosphor screen for re-conversion of photons. Photons first enter the intensifier input window and strike the photocathode where they are converted to photoelectrons. When the voltage of the photocathode is negative relative to

the MCP entrance face, the photoelectrons are pulled toward the MCP and the gate is in the 'on' state. When the polarity is reversed, photoelectrons remain at the photocathode and the gate is in the 'off' state. photoelectrons at the MCP are accelerated along a series of thin (10 μm) channels where collision with the inner walls produce secondary electrons. At the output of the MCP, the electrons are drawn to a phosphor screen which is held at a positive voltage relative to the MCP output. Electrons colliding with the phosphor screen produce green photons. A lens is used to relay photons from the intensifier to the CCD where they are converted to photoelectrons. During the exposure time of the CCD, charge wells are filled with photoelectrons which are subsequently read out at the end of the exposure. To collect the full temporal response to a laser pulse, the intensifier gate is in the 'on' state at different time delays relative to the pulse of the laser.

The specifications of our ICCD system is as follows. A 12-bit cooled CCD camera (Picostar HR-12 CAM 2, LaVision GmbH Goettingen Germany) with resolution of 1376 $\times$ 1024 pixels is mounted to a voltage-gated image intensifier (Picostar HR-12, LaVision GmbH, Goettin- gen, Germany). The system is synchronized relative to the 80 MHz pulse repetition rate of the Ti-Sapph laser through a constant fraction discriminator (Becker & Hickl GmbH) using either the electronic trigger of the laser or directly from the optical pulses. A fast HRI delay unit was used to set the gate opening at different times relative to the laser pulse. Typical CCD integration times of 100 ms corresponds to data acquisition at 8 million laser pulses for each image. To project the image onto the photocathode of the image intensifier, a camera lens (AF Nikkor, f2.8 Nikon) is mounted to the intensifier. For fluorescence measurements, 2-inch emission filters were attached to the lens.

3.3 Experimental Measurement and Reconstruction Procedures

In this section, we summarize the steps taken when performing a standard fluorescence tomography experiment. We divide the procedure into two parts, the acquisition of data and the processing of data for reconstruction.

3.3.1 Measurement Procedure

1. **Sample Preparation**: The sample is placed on the imaging platform and positioned such that all fiducial markers are within the CCD field of view. For in vivo measurements anesthesia is delivered through tubes attached to the imaging plate.

2. **Excitation Measurement**: A measurement at the excitation wavelength of the fluorophore is obtained through the sample. This data can be used to determine the relative transmission of the excitation light at different sources or for full 3D reconstruction of optical properties throughout the medium.

3. **Calibration Measurement**: An excitation measurement is obtained through a calibration phantom with known optical properties (standard epoxy resin phantom, $\mu_a = 0.1\ \mathrm{cm}^{-1}$, $\mu_s' = 10\ \mathrm{cm}^{-1}$). The scaled measurement data, y_{scaled} can be written in terms of the actual measurement data, y_{meas}, measured calibration data, $y_{meas-calib}$ and simulated calibration data, $y_{sim-calib}$:

$$y_{scaled} = y_{meas} \frac{y_{sim-calib}}{y_{meas-calib}} \tag{3.1}$$

   A calibration measurement allows both the cancellation of source and detector coupling coefficients and the conversion of the units of the experimental data (CCD counts) to the physical units of the Monte Carlo simulation.

4. **Fluorescence Measurement**: The appropriate emission filter is placed in front of the lens. A fluorescence measurement is obtained through the sample of interest.

5. **Surface Capture Measurement**: For capturing the 3D surface boundary using the photo-stitching technique, multiple white-light images are taken with a standard camera centered on the sample. The camera is rotated around the sample and positioned at various angles relative to the sample.

6. **IRF Measurement**: The sample is removed from the imaging platform and replaced with a white paper. The instrument response function (IRF) is measured through the paper at every source and detector location. In many measurements, detectors are taken at the same x-y coordinates as the sources.

3.4 Reconstruction Procedure

1. **Surface Reconstruction**: The software Autodesk 123D Catch is used to reconstruct the 3D scene from the white light images. The output of Autodesk 123D Catch is the photo scene in the form of a 3D surface mesh (.obj format). Using MeshLab, extraneous elements of the scene are removed. The mesh is imported into Matlab for alignment and then converted into a volumetric binary cloud used for forward modeling. Fig. 3.2(b) shows the reconstructed surface of a nude mouse under anesthesia after processing from the original captured scene in Fig. 3.2(a). Fig. 3.2(c) shows surface reconstruction can be successfully applied to smaller samples such as excised mouse organs which can be used for ex vivo optical tomography measurements.

2. **Co-registration**: The affine transformation which maps the coordinates of the CCD camera to the reconstructed 3D surface is generated using fiducials (CT spots) which are visible

from the CCD white light image and on the surface mesh.

3. **Source and Detector Positions**: The source positions are determined by first measuring the IRF for each source and then by finding the pixel corresponding to the maximum intensity for each measurement. Detectors are chosen as subsets of CCD pixels within the coverage area of the sources. Sources and detectors are mapped into the surface mesh coordinate system using the affine transformation from the previous step.

4. **Monte Carlo Simulation**: The source and detector positions and the volumetric binary cloud are used as inputs into the time-domain Monte Carlo software, tMCimg (Boas et al., 2002). The output of the Monte Carlo simulation are the source and detector Green's functions. These are also frequently referred to as the two-point functions.

5. **Weight Matrix Calculation**: The weight matrix (or sensitivity matrix) is generated using the source and detector's Green's functions. Since the shape of the IRF was found to depend on detector location (section 3.5.2), the weight matrix for each detector was convolved with the experimentally measured IRF at each detector.

6. **Reconstruction Algorithm**: Linear inversion is performed using the SVD implementation of Tikhonov regularization.

3.5 System Characterization and Calibration

3.5.1 Flat Field Correction

The flat field correction technique can compensate for differences in sensitivity over the field of an image. This sensitivity difference, due to both the sensitivity variations in the intensifier phosphor and CCD pixel sensitivity can be measured and used to correct measurement data. Two images

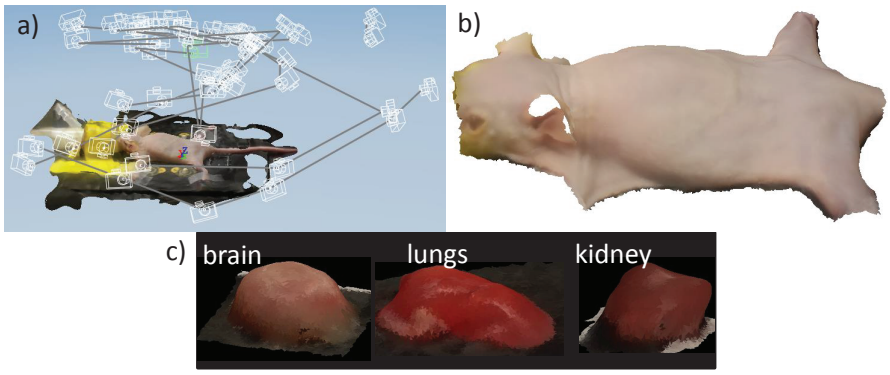

Figure 3.2: Demonstration of the photo-stitching technique applied to an anesthetized mouse and excised mouse organs. (a) 3D scene of anesthetized mouse with camera positions for each white light image displayed. (b) 3D surface of mouse after a processing step. (c) Capture 3D surface of mouse organs which were used for ex vivo imaging.

are required for this technique, a blank image, I_{blank} under an uniform light source and a dark image, I_{dark} with the camera shutter closed. The corrected image, $I_{corrected}$ in terms of the measured image, I is given by:

$$I_{corrected} = \frac{I - I_{dark}}{I_{blank} - I_{dark}} \tag{3.2}$$

Both I_{blank} and I_{dark} are dependent on system parameters such as intensifier gatewidth settings. An example of a blank image used for flat field correction is shown in Fig. 3.3. The image was generated by averaging 50 measurements of a uniform light source for intensifier gain = 500 V and gatewidth = 500 ps.

3.5.2 Instrument Response Function

The instrument response function accounts for multiple experimental parameters such as the pulse width of the laser, the shape of the gate of the image intensifier and the broadening of the laser pulse inside the optical fiber. It is used to estimate the time t_0 that the laser is first incident into the surface of the imaging medium. As part of our experimental procedure, we measure the IRF using excitation light focused onto a piece of white paper. The measured IRF is directly convolved

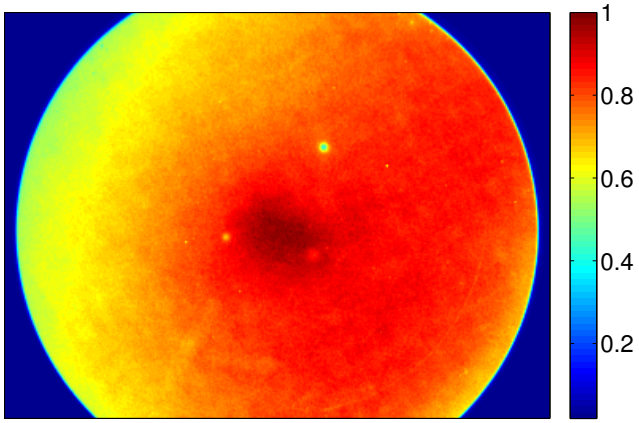

Figure 3.3: Example blank image used for flat field correction for intensifier gain = 500 V an gatewidth = 500 ps.

with the time domain weight matrices before image reconstruction. We have observed a dependence of the measured IRF shape on the spatial location of the source. This effect is shown in Fig. 3.4 where the IRF's measured for sources along a line in either the x or y direction are plotted. To account for this dependence, the IRF for every detector must be measured. Also in the reconstruction algorithm, the temporal profile of the weight matrix for a particular source/detector pair and voxel is convolved with the IRF for the corresponding detector.

3.5.3 Timing Jitter

An important characterization of a time domain system is the stability of the IRF over time. We measured the timing jitter of our system over 45 minutes, which is the typical measurement time for a combined excitation and fluorescence measurement of a mouse. The jitter at each time point was obtained by fitting for the time shift Δt which produced the best match between the measured IRF and a reference IRF at $t = 0$. It can be seen that the mean of the jitter stays close to zero over the duration of this measurement while maximum absolute jitter did not exceed 20 ps (Fig 3.5). However, in separate measurements we have observed two effects which produce more

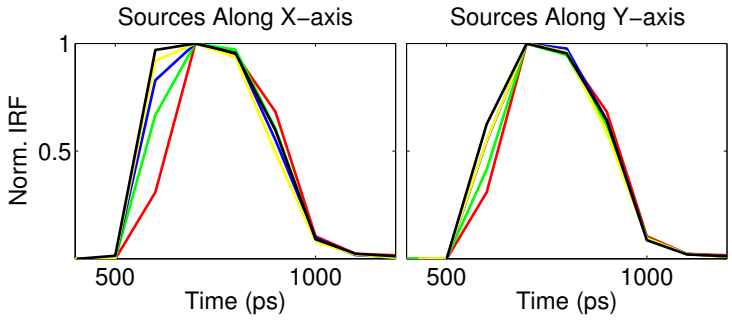

Figure 3.4: The dependence of the IRF on the location of a point source. The IRF's for five sources scanned in a line along the x-direction (left) and y-direction (right) are shown with different colors.

significant changes in the IRF. First, the shape of the IRF can abruptly change along with a significant shift in t_0 by as much as 50-100 ps. Second, the mean timing jitter can undergo gradual drifting which can exceed 50 ps. This drift occurs more frequently in longer experiments (> 2 hours). Since accurate optical property reconstructions are particularly sensitive to timing jitter, a method was devised to reduce overall jitter.

To track the amount of timing jitter, a reference fiber source was incorporated into our system using a 1×2 fiber coupler. The IRF was obtained simultaneously with each measurement. If the jitter between a reference IRF and the measured IRF is above a preset threshold, the measurement is reacquired. This reduces the overall jitter in the acquired data set at the expense of additional acquisition time.

3.6 Dynamic Power Control

It is known there exist large variations in light transmission through the mouse body due to differences in tissue absorption and mouse thickness. This precludes obtaining a high signal to noise ratio for all regions of the mouse. This issue is illustrated in Fig 3.6(a) showing the transmission through a live mouse with a lung tumor model consisting of iRFP720 cells. The high signal from the lateral chest region approaches the dynamic range limit of the CCD camera while the central

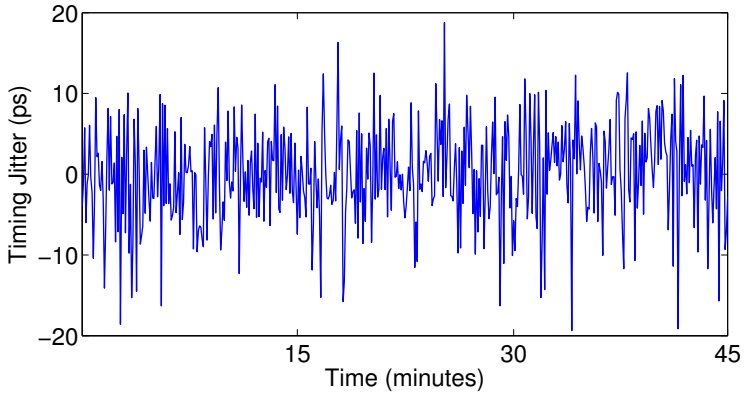

Figure 3.5: Stability measurement for the IRF. The amount of timing jitter in the IRF is quantified for a single 45 minute measurement.

chest area has very low SNR. We address this problem in two ways: dynamic laser power control and spatial filters.

The laser power reaching the mouse can be dynamically controlled using a filter wheel containing neutral-density (ND) filters (0.1 to 3 OD). At each source position in a tomography measurement, a preliminary scan is performed for a range of time gates close to the peak of the TPSF. Appropriate ND filters are chosen based on this initial scan and the TPSF is taken with these settings. The measurement data is scaled based on the source dependent ND filters so that reconstruction is performed on data with the same effective input laser power. This method is able to produce uniformity in the maximum SNR for different sources in a scan. However, a large spatial variation in SNR can still remain for a single source.

To account for this problem, we have also incorporated a spatial filter technique to achieve high SNR data throughout the mouse. In this technique, areas of high transmission are completely blocked from the camera. This allows the signal for low SNR regions to be increased either through an increase in laser power or through longer exposure time of the camera. The filtering is implemented using a mechanical slide whose size is adjustable in both axes. The filter is placed $\sim$5 cm above the mouse surface. To demonstrate the utility of a spatial filter, we compare data

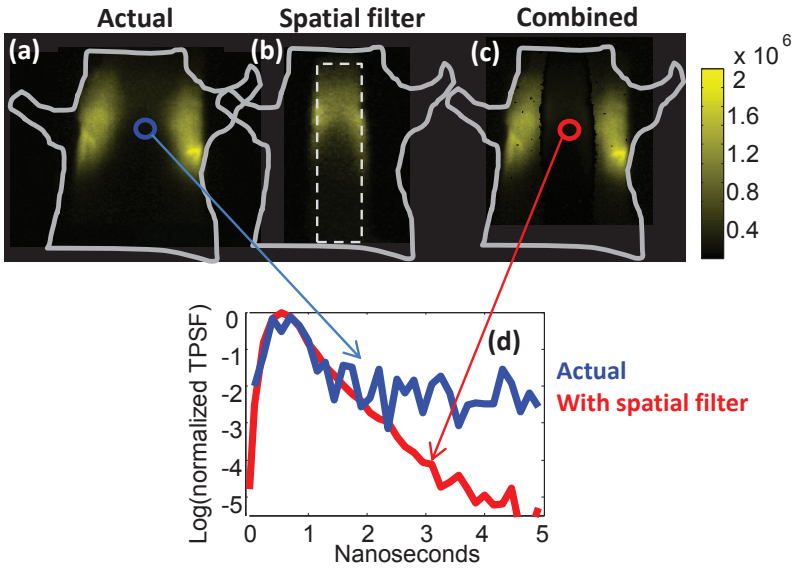

Figure 3.6: Improved SNR performance using spatial filers. (a) Light transmission through the chest of a mouse with near-infrared fluorescent proteins in the lungs. (b) Light transmission when lateral chest region is masked out using spatial filter. (c) Combination of the transmission data from (a) and (b) while scaling for input laser power.

acquired without a spatial filter (Fig. 3.6(a)) with data acquired with a spatial filter over the lateral chest region (Fig. 3.6(b)). While combining the two data sets with proper scaling produces an intensity image (Fig. 3.5(c) with similar appearance to Fig. 3.6(a), it is clear from the TPSF's taken from the central chest region where the transmission of light is low that the exponential decay of the iRFP720 fluorescence is measured much more accurately with a spatial filter.

3.7 Acquisition Software

A custom Matlab GUI was written for data acquisition to replace a previous proprietary acquisition software. Serial communication is initiated to various hardware devices (CCD camera, image intensifier, delay box, translation stage and filter wheel). The laser trigger from the constant fraction discriminator (CFD) is automatically detected. Coregistration between the x-y translation stage coordinate system and the CCD camera field of view is implemented using a calibration

procedure at four points on the stage. This allows for easy selection of source locations and manual positioning of the laser source relative to the sample. Different modes of operation include single or continuous scanning for a set of time gates and sources. The software includes the aforementioned methods for higher quality data acquisition such as dynamic power control, which allows a preliminary scan to determine appropriate source position dependent input laser power and IRF jitter control which discards measurements containing high timing jitter.

4

Time Domain Fluorescence Reconstruction in the Asymptotic Regime

In this chapter, we will focus on image reconstruction using the decay portion of the fluorescence temporal point spread function which contains information about the intrinsic fluorescence lifetimes of fluorophores within a turbid medium. We will consider the case of reconstruction of a single lifetime component but more emphasis will be placed on multiplexing multiple fluo-

rophores using lifetime as contrast. Although fluorescence lifetime multiplexing is being widely used in fluorescence lifetime imaging microscopy (FLIM) (Bastiaens & Squire, 1999; Selvin, 2000; Berezovska et al., 2003), a technique applicable for tissue sections, we and others look to extend this concept to tomographic separation of multiple fluorophores with the ultimate goal of visualization of multiple molecular processes simultaneously in the entire mouse.

Traditionally, a common method to perform time domain fluorescence tomography reconstruction was by taking TD data for a set of time gates and stacking them for every source/detector pair into a measurement vector. Standard linear inverse techniques were used to invert the measurement vector with the full time domain weight matrix (Chen et al., 2011; Holt et al., 2012). We refer to this approach as the Direct TD (DTD) approach. In (Kumar et al., 2005, 2006), an asymptotic TD (ATD) approach for lifetime multiplexing was introduced where reconstruction was performed in two steps. First, a multiexponential fit is performed on late time gates to recover decay amplitudes. Second, the decay amplitudes for each fluorophore are independently inverted using CW weight functions. It was shown with simulations that ATD could separate inclusions which were 7 mm, 5 mm and 3 mm apart while a DTD approach using early time gates produced reconstructions with significant cross-talk and inclusions could not be correctly localized. A separate study (Kumar et al., 2008a) comparing lifetime multiplexing using an ATD approach with a frequency domain (FD) approach applied to Fourier transformed TD data showed that the ATD approach could accurately localize the positions of inclusions 3 mm apart in a simulation and 6 mm apart in an experimental phantom while reconstructions using FD data could separate neither.

We build on this previous work in the following ways. We show experimentally the separation of fluorescent inclusions down to 1.4 mm using the ATD approach to validate previous simulation results. We express the ATD approach in matrix form by decomposing the TD weight matrix for late time gates as a product of a temporal basis function matrix and a spatial block diag-

onal matrix. We compare the linear inverse operators for the ATD approach and DTD approach and show they are mathematically distinct. We show that within a Bayesian framework, both the DTD and ATD approaches can produce inverse operators that can be applied either directly to the TD data or to the decay amplitudes and ATD can be obtained from DTD through a substitution of the measurement covariance matrix. We introduce a quantitative measure for cross-talk and show that the ATD approach provides reconstructions with superior cross-talk performance and better separation than DTD. We show through simulation that ATD can provide better estimates for the relative amounts of two fluorophores than DTD. Finally, we revisit the asymptotic condition and explore the relationship between the diffuse temporal response and the intrinsic fluorescence lifetime. We introduce a new time constant for determination of the beginning of the asymptotic regime. Finally, we plot contours for the diffuse time constant and recovered fluorescence lifetime as a function of background optical properties using simulation. Prior to introducing the mathematical form of the time domain algorithms, we describe a source of error inherent to multiplexing experiments known as cross-talk.

4.1 Cross-Talk for Lifetime Multiplexing

When multiplexing multiple parameters, an important performance measure, along with resolution is inter-parameter cross-talk. Cross-talk can be understood as the interference caused by the presence of one type of parameter on the reconstruction of other types of parameters. This is typically illustrated in the two parameter case by considering a medium where objects of the two parameter types are separated spatially. When image reconstruction is performed, regions which should only reconstruct for one type of object will show reconstructed values for both types. In the case of closely spaced objects, cross-talk has also been shown to result in localization error (Kumar et al., 2006). In certain optical tomography problems, cross-talk is known to be high due to

non-uniqueness of the inverse problem. In the field of diffuse optical tomography (DOT), it was shown that multiple absorption and scattering distributions can produce the same single wavelength CW measurement at the surface (Arridge & Lionheart, 1998). Reconstruction with CW data results in indistinguishable reconstructions for absorption and scattering inclusions. Similarly in fluorescence tomography, since lifetime information cannot be obtained from CW data, reconstruction of multiple lifetimes with CW data results in a large cross-talk component. In later sections of this chapter, we will develop an intuitive method to quantify the cross-talk performance of an experimental technique or reconstruction algorithm for multiplexing which to our knowledge has not previously been applied.

4.2 Time Domain Fluorescence Tomography Forward and Inverse Problems

Consider a diffuse medium embedded with N fluorophores, characterized by yield distributions $\eta_n(\mathbf{r})$ and lifetimes τ_n. The fluorescence temporal point spread function (TPSF) for impulsive excitation at source position $\mathbf{r}_s$ and detection at $\mathbf{r}_d$ on the surface of a bounded transport medium of support Ω takes the form:

$$U_f(\mathbf{r}_s, \mathbf{r}_d, t) = \sum_{n=1}^{N} \int_{\Omega} W_n(\mathbf{r}_s, \mathbf{r}_d, \mathbf{r}, t)\eta_n(\mathbf{r})d^3r, \tag{4.1}$$

where $W_n = G^x(\mathbf{r}_s, \mathbf{r}, t) \bigotimes \exp(-t/\tau_n) \bigotimes G^m(\mathbf{r}, \mathbf{r}_d, t)$ is the TD sensitivity function, expressed as a double convolution of the source and detector Green's functions, G^x and G^m, for light transport in the medium (Arridge, 1999), and a fluorescence decay term. Discretizing the medium volume into V voxels, (4.1) can be written as a linear matrix equation for M source-detector pairs, L time

gates and N lifetime components:

$$\begin{bmatrix} y(t_1) \\ \vdots \\ y(t_L) \end{bmatrix} = \begin{bmatrix} W_1(t_1) & \dots & W_N(t_1) \\ \vdots & \ddots & \vdots \\ W_1(t_L) & \dots & W_N(t_L) \end{bmatrix} \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_N \end{bmatrix} \tag{4.2}$$

$$y = W\eta, \tag{4.3}$$

where y is a $(ML \times 1)$ vector of the measured fluorescence data U_f, $W = [W_1, ..., W_N]$ is a $(ML \times VN)$ matrix of TD sensitivity functions and $\eta = [\eta_1, ..., \eta_N]^T$ is a $VN \times 1$ vector of yield distributions for all lifetimes. A direct inversion of (4.3) leads to the direct TD (DTD) approach, and the inverse problem is expressed in the standard way using Tikhonov regularization (Bertero & Boccacci, 1985) to recover η:

$$\eta^{DTD} = W^T(WW^T + \lambda I)^{-1}y. \tag{4.4}$$

It is possible to recast the TD forward problem in an alternate form in the asymptotic limit, i.e., for times much longer than the intrinsic diffuse timescale τ_D of the medium, and assuming the widely held condition, $\tau_n > \tau_D$ (Kumar et al., 2005). The $W_n(t)$'s in (4.1) can then be factored into a product of CW sensitivity functions and simple exponential decays (Kumar et al., 2005, 2006), for arbitrary transport media.

While previously, this separation of the temporal and spatial components was used in a two step reconstruction algorithm without consideration for the underlying matrix operations, we show additional insights can be gained when the forward problem in (4.2) is first expressed as a product of two matrices:

$$y \overset{t \gg \tau_D}{=} A\overline{W}\eta, \tag{4.5}$$

where $A = [\exp(-t/\tau_1) * I, ..., \exp(-t/\tau_N) * I]$ is a ($ML \times MN$) basis matrix of exponential decays, I is a (M $\times$ M) identity matrix, and $\overline{W} = diag(\overline{W}_n)$ is a ($MN \times VN$) block diagonal matrix containing CW weight matrices, $\overline{W}_n$, which are evaluated with the background absorption reduced by $1/v\tau_n$.

A simple manipulation of (4.5) can allow a fundamentally different approach to solving the TD inverse problem in the asymptotic limit. Since A is a well-conditioned matrix, it can be first inverted without regularization by multiplication with its Moore-Penrose pseudoinverse $A^\dagger$. The left hand side of (4.5), after pre-multiplication with $A^\dagger$, is then simply a linear least squares solution for a multi-exponential analysis of U_f:

$$a = \overline{W}\eta \tag{4.6}$$

where $a = [a_1, ..., a_N]^T$ is a ($MN \times 1$) vector of decay amplitudes for all source/detector pairs and lifetimes. In the next step, we apply Tikhonov regularization to invert the decay amplitudes:

$$\eta^{ATD} = \overline{W}^T(\overline{W}\overline{W}^T + \lambda I)^{-1}A^\dagger y. \tag{4.7}$$

Application of the above inverse matrix operator is equivalent to the two step asymptotic TD (ATD) approach from (Kumar et al., 2008b). When both DTD and ATD are expressed in their matrix forms, it is clear that although their forward problems ((4.3) and (4.5)) are the same for late time gates, the corresponding inverse problems ((4.4) and (4.7)) are distinct, and will produce different reconstructions even when applied to the same measurement. The key aspect of the ATD approach is that the inversion of the basis matrix A is removed from the regularized inversion. This ensures that the measurements in y are directly separated using the exponential basis function of each fluorophore. Additionally, given the block diagonal nature of $\overline{W}$, (4.7) essentially reduces

to completely separate inverse problems for each yield distribution η_n. This should be contrasted with the DTD inversion in (4.4), where the inverse problem for each η_n is not separable. Consequently, the ATD approach results in significantly lower cross-talk between the yields of multiple lifetimes than the DTD approach, as we will demonstrate using simulations and experiments later in the chapter.

4.3 Bayesian Interpretation

In the previous section, we have shown that the DTD and ATD methods produce distinct linear inverse operators for the TD fluorescence tomography problem. In this section, we show that both DTD and ATD can be expressed either in terms of the full TD weight matrix W or the block diagonal weight matrix containing the CW weight matrices, $\overline{W}$. We use a Bayesian framework to interpret the dual forms of each method.

We start with the ATD inverse operator from (4.7) and derive its alternate form as follows:

$$\begin{aligned} W_{ATD}^{-1} &= \overline{W}^T(\overline{W}\overline{W}^T + \lambda I)^{-1}A^\dagger \\ &= \overline{W}^T A^T (A^T)^\dagger (\overline{W}\overline{W}^T + \lambda I)^\dagger A^\dagger \\ &= \overline{W}^T A^T (A\overline{W}\overline{W}^T A^T + \lambda A A^T)^\dagger \qquad (4.8) \\ &= W^T (WW^T + \lambda A A^T)^\dagger \qquad (4.9) \end{aligned}$$

where step 3 results from application of the generalized inverse product rule $(ABC)^\dagger = C^\dagger B^\dagger A^\dagger$ (Hartwig, 1986). Similarly, the DTD inverse operator (from (4.4)) can also be written in terms of

the block diagonal weight matrix, $\overline{W}$:

$$W_{DTD}^{-1} = W^T(WW^T + \lambda I)^{-1} \tag{4.10}$$

$$= \overline{W}^T(\overline{W}\overline{W}^T + \lambda(A^TA)^{-1})^{-1}A^\dagger \tag{4.11}$$

Since the DTD and ATD solutions can be obtained using either form, implementation considerations should dictate which form is used. We will show in section 4.6 that the $\overline{W}$ form applied to amplitudes has significant advantages in terms of computation time and memory storage over the W form. We next interpret the results in (4.9) and (4.11) using a Bayesian approach. First we assume that the fluorescence yield, η and the noise corrupting the measurement are drawn from white Gaussian distributions with covariance matrices $C_\eta = \sigma_\eta^2 I$ and $C_y = \sigma_y^2 I$ respectively. For $\lambda = (\sigma_y/\sigma_\eta)^2$, the DTD inverse operator (4.9) is the minimum mean square estimator (MMSE) for η (Bertero & Boccacci, 1985). The ATD solution can be understood in two different ways with respect to the DTD solution. From (4.8-4.11), it is clear that the DTD inverse operator is identical to the ATD inverse operator under two conditions. (i) When the amplitude covariance matrix $(A^TA)^{-1}$ is replaced by a diagonal matrix λI in (4.9). (ii) When a measurement covariance AA^T is added within the DTD inverse operator in (4.11).

Given accurate assumptions about the noise model and for appropriate choice of λ, the DTD solution will have lower mean square error than solutions from all other estimators including the ATD solution. However, we have shown in the previous section that ATD reconstructions may perform better using other measures such as reduced cross-talk and improved localization of closely separated inclusions. In the next section, we will attempt to gain insight into why ATD approach has better cross-talk performance than the DTD approach and how this improvement is generally true for all voxels in an arbitrary diffuse medium.

4.4 Resolution Matrix

A useful tool for analyzing the properties of an inverse operator is its resolution matrix. It is defined as the product of the inverse operator, W^{-1} and the forward matrix, W:

$$R = W^{-1}W \tag{4.12}$$

In the case of reconstructing yield for a single lifetime, the j^{th} column of R represents the point spread function due to an inclusion at the j^{th} voxel in the imaging medium. In the case of reconstructing the yields at N lifetimes and for V voxels in the imaging medium, the $(n-1)V + j^{th}$ column of R contains a general response due to an inclusion at the j^{th} voxel and for the n^{th} lifetime. This response includes both the point spread function for lifetime n and the point spread function for all other lifetimes. This leads to a natural definition of cross-talk for a voxel j of lifetime n as the magnitude of the point spread function for j at lifetimes different from n. In Fig. 4.1, we show the form of the resolution matrix in the case of two lifetimes. The cross-talk for every voxel in the imaging medium can be determined from the off diagonal blocks of the resolution matrix.

For comparison of the cross-talk for the DTD and ATD methods, we derive the form of their respective resolution matrices:

$$R_{DTD} = W_{DTD}^{-1}W \;\; = \;\; \overline{W}^T(\overline{W}\overline{W}^T + \lambda(A^TA)^{-1})^{-1}A^\dagger(A\overline{W}) \tag{4.13}$$

$$R_{ATD} = W_{ATD}^{-1}W \;\; = \;\; \overline{W}^T(\overline{W}\overline{W}^T + \lambda I)^{-1}A^\dagger(A\overline{W}) \tag{4.14}$$

It can be observed that since $A^\dagger A = I$ (a property that holds when A is full column rank (Ben-

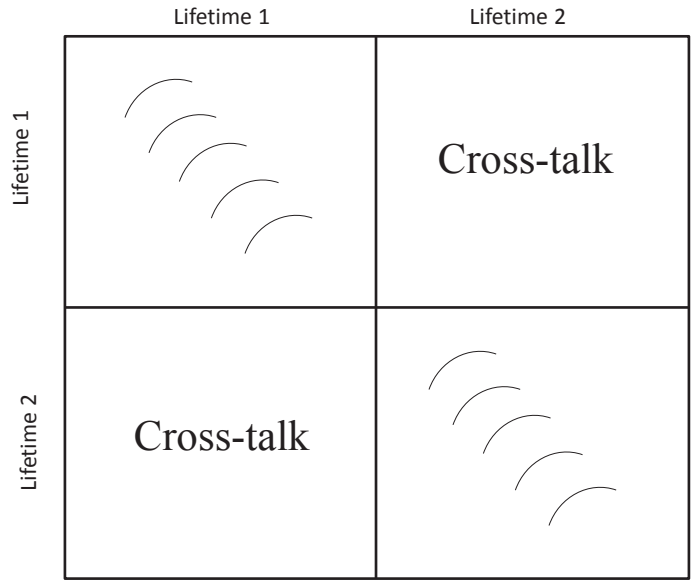

Figure 4.1: The form of the resolution matrix for multi-parameter imaging problems. The on-diagonal blocks represent the individual parameter responses. The off-diagonal blocks represent cross-talk between two parameters.

Israel & Greville, 2003)) and every remaining term in (4.14) is block diagonal, the resulting R_{ATD} is also a block diagonal matrix. Therefore, ATD reconstructions have zero cross-talk. On the other hand, R_{DTD} is generally not block diagonal and hence DTD reconstructions contain non zero amounts of cross-talk. This observation provides justification for replacement of the measurement covariance matrices in the ATD approach from the last section. Essentially, in ATD, the measurement covariance matrix which leads to the minimum error is replaced by one which leads to zero cross-talk. We will verify the forms of the resolution matrix in simulations in the next section.

4.5 Simulation Results

We next quantified the above theoretical results using Monte Carlo simulations, performed using tMCimg (Boas et al., 2002). The simulation medium was a 2 cm $\times$ 2 cm $\times$ 2 cm slab with optical properties of $\mu_a = 0.2\,\text{cm}^{-1}$ and $\mu_s' = 10\,\text{cm}^{-1}$. 49 equally spaced sources and detectors were

located at $z = 0$ cm and $z = 2$ cm planes respectively. The time step was set to 100 ps. Fluorescent inclusions (0.125 mm^3) were placed at the center of the slab. A 2% shot noise was added to all forward data prior to reconstruction. In all simulations, the regularization parameter λ was chosen corresponding to the least reconstruction error ($E = \|\eta_{recon} - \eta_{true}\|$).

4.5.1 Single Lifetime Reconstruction

Figure 4.2 shows simulation results for a single fluorescent inclusion using the DTD approach. The inset of Fig. 4.2(a) shows that the resolution, quantified by the full volume at half maximum (FVHM) of the reconstructed yield, is poorer with later time gates. Also, a singular value decomposition (SVD) analysis of the DTD sensitivity matrix was performed. SVD analysis has been used previously in optical tomography to explore optimizing experimental parameters such as source/detector number and coverage and voxel size (Graves et al., 2004). The slope of the singular value spectrum is an indicator for the conditioning of the inverse problem. We see from Fig. 4.2(b) and 4.2(c) that SVD analysis shows distinct behavior for the early and late portions of the TPSF. The slope of the singular value spectra increased with the number of early time gates, indicating improvement in the conditioning of the inversion and gains in resolution (Fig. 4.2(b)). However, the SVD spectra show negligible change for any combination of multiple late time gates (Fig. 4.2(c)). This demonstrates the redundancy of using multiple gates in the asymptotic region for tomography, as we expect from the spatio-temporal factorization in (4.5).

4.5.2 Lifetime Multiplexing

We next compared the imaging performance of ATD and DTD methods using two inclusions with distinct lifetimes ($\tau = 0.87$ ns and 1.27 ns) and with varying separations of 2 mm to 6 mm. Figure 4.3 shows a comparison of ATD with DTD applied to the same set of 12 late time gates.

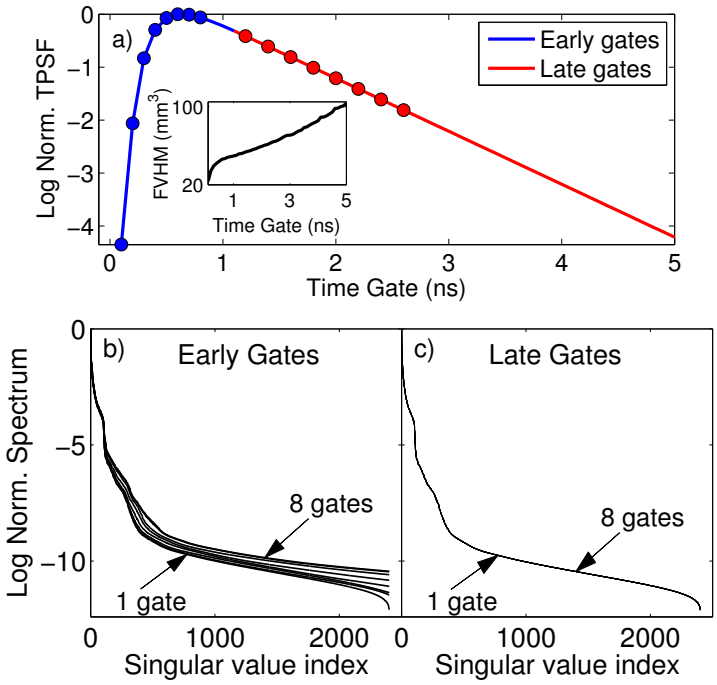

Figure 4.2: (a) Fluorescence TSPF for a 2 cm thick diffusive slab with a fluorescent inclusion ($\tau = 1$ ns) at the center. The inset shows the FVHM of the reconstructed yield for individual time gates. SVD spectra of the DTD weight matrix are shown for (b) one to eight early gates [indicated as blue circles in (a)] and (c) one to eight late gates [red circles in (a)]. The gates were stacked in order of decreasing intensity for both cases.

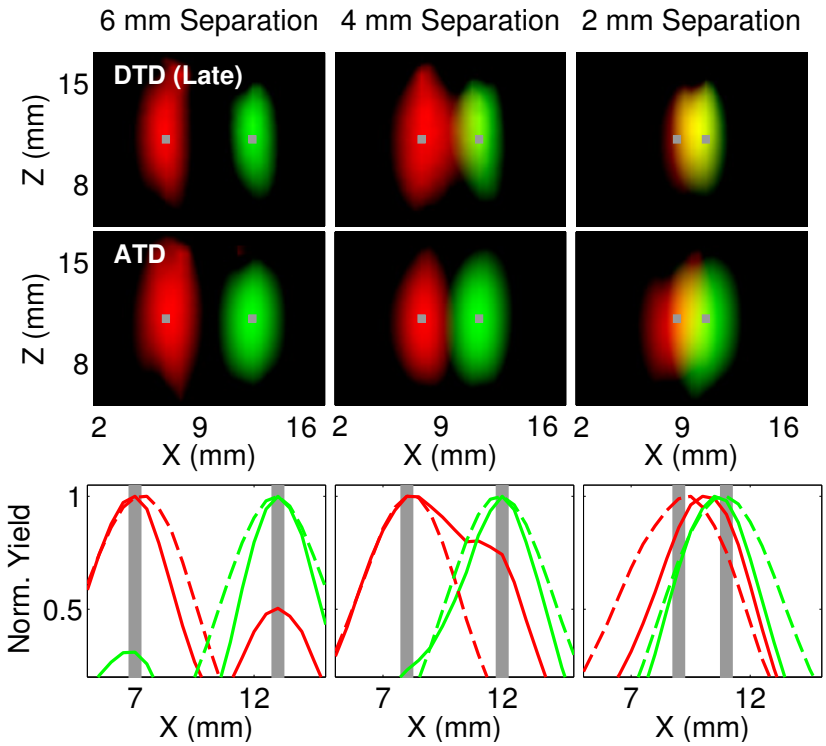

Figure 4.3: Comparison of reconstructions obtained from applying ATD and DTD to the same set of 12 late time gates. Red and green corresponds to the yield distributions for 0.87 and (1.27 ns), respectively. The true locations of the inclusions are shown in gray. In the first two rows, the X–Z plots are generated by assigning the recovered yields to the red (0.87 ns) and green (1.27 ns) components of the RGB colormap. Each distribution is thresholded at 50% of its maximum. The bottom row shows line plots for ATD (dashed line) and DTD (solid line) along the x axis at the depth of the inclusion. The computation time for ATD was 21 times shorter than that for DTD.

It can be seen from the line plots that there is significantly higher cross-talk for reconstructions using DTD (Fig. 4.3, yellow indicates cross-talk), which precludes accurate localization of the inclusions even for 4 mm separation. The ATD approach, however, provides minimal cross-talk, and correctly localizes the inclusions down to 2 mm separation.

The resolution matrices for DTD and ATD, R_{DTD} and R_{ATD} were computed using (4.13) and (4.14). To aid with visualization, the rows and columns of the matrices were binned down by a factor of 80 and normalized to the maximum value of the matrix. Fig. 4.4 shows plots of the resolution matrices for both methods. It can be seen that R_{DTD} contained non-zero terms in the off

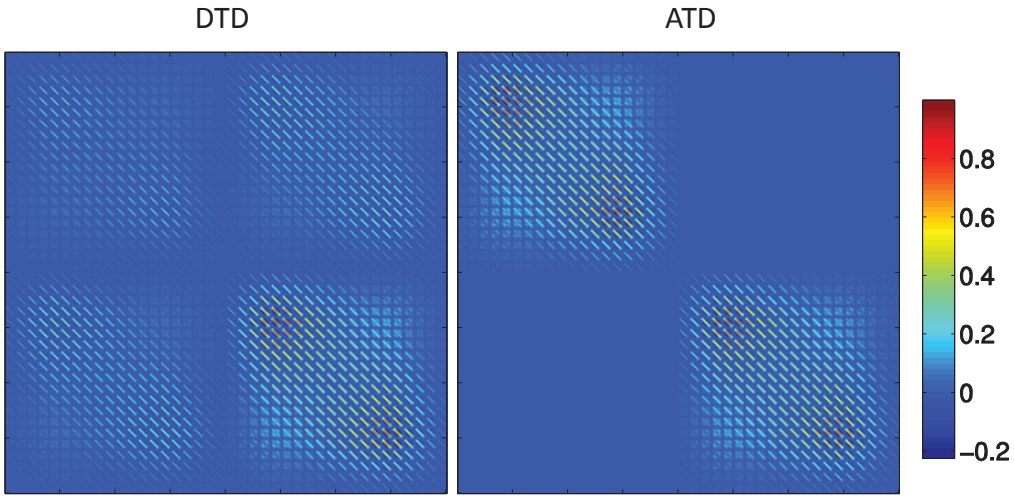

Figure 4.4: The resolution matrices for DTD and ATD. Matrices are binned down by factor of 80 and normalized to their maximum values.

diagonal block while R_{ATD} was exactly zero for all off diagonal terms. This indicates that cross-talk is minimized for ATD reconstructions compared to DTD reconstructions over all voxels of the imaging medium. Also, R_{DTD} showed a significant difference in magnitude for the two lifetime components on the main diagonal. This difference could lead to inaccuracy when determining relative yields of the two lifetimes. R_{ATD} showed a much closer relative magnitude for the two lifetime components on the diagonal blocks.

To demonstrate the quantitative accuracy of the ATD method in estimating the relative amounts of two fluorophores, we performed a simulation with two fluorescent inclusions of differing lifetimes ($\tau = 0.87$ ns and 1.27 ns) placed at the same location in the slab at the z = 1cm depth. The ratio of the fluorescence yields of the inclusions were varied from 1:1 to 1:6. Both ATD and DTD reconstructions were performed on noise corrupted measurements. To calculate the standard deviation of the estimates, measurements formed from 50 different realizations of the noise were reconstructed. In Fig. 4.5, the recovered yield using DTD (blue) and ATD (red) are shown with the true ratio in black and the standard deviation as error bars. The ATD method is able to recover

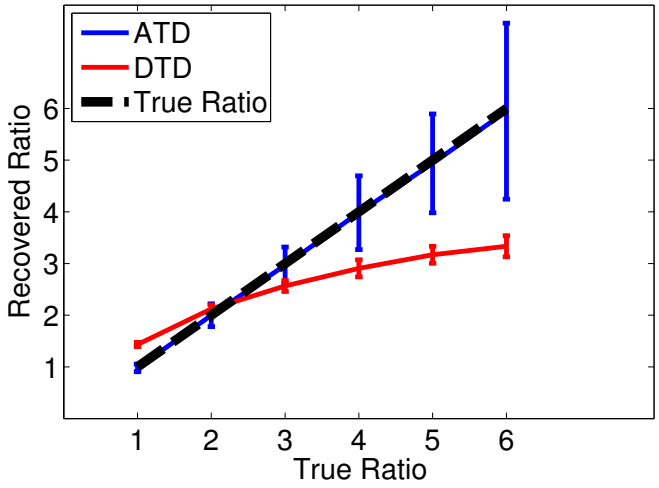

Figure 4.5: Determination of relative amounts of two fluorophores using ATD and DTD. The recovered yield using DTD (blue) and ATD (red) are shown with the true ratio in black.

the true yield ratios while, due to cross-talk, the DTD estimates are inaccurate. It can be observed for high true ratio values, the standard deviation for the reconstructed ratio was higher for ATD than DTD. This can be attributed to the fact that for large ratios, the uncertainty in the smaller lifetime component is amplified. Since the ATD method was able to accurately reconstruct large ratios while the DTD method underestimated large ratios, the uncertainty in the smaller lifetime component was greater for ATD than DTD resulting in higher uncertainty in the recovered ATD ratios.

4.6 Experimental Results

Finally, we validated our simulation results using phantom experiments. We constructed three phantoms consisting of cell culture dishes [88 mm diameter, 17 mm in height, with two parallel tubes (0.965 mm outer diameter, 0.58 mm inner diameter) spanning the length of the dish, at a height of 8 mm, and a center-to-center separation of 5.6, 2, or 1.4 mm (see Fig. 4.6)]. The tubes were filled with 3, 30-diethylthiatricarbocyanine iodide (DTTC) dye at a concentration of either

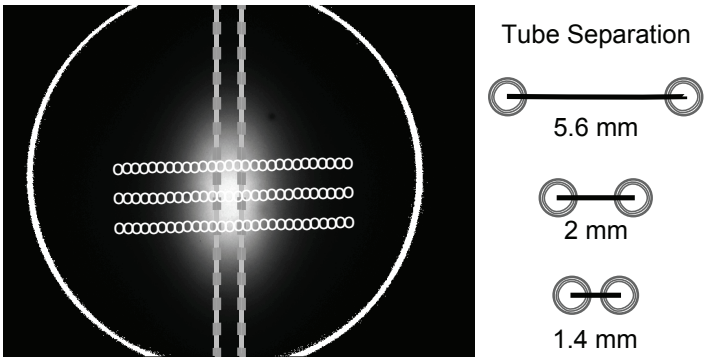

Figure 4.6: Dish phantom filled with Intralipid and nigrosin solution to a depth of 1.7 cm, shown with the tubes. Parallel dye filled tubes have a center-to-center separation of 5.6 (shown), 2, or 1.4 mm. Also shown are source and detector positions (o) overlaid on top of the total CW fluorescence intensity distribution for all sources.

7.4 µM in 100% ethanol ($\tau_L = 1.27$ ns) or 9.2 µM in 66% ethanol, 33% H_2O ($\tau_S = 0.88$ ns). These concentrations were selected to achieve similar CW fluorescence yield for the two dyes in the absence of a scattering medium. The fluorescence lifetime of each DTTC solution was measured directly in the absence of a scattering medium. All phantoms were filled with 99 mL of 1% Intralipid with 4.34 µg / mL of nigrosin, resulting in a reduced scattering coefficient of 10 cm^{-1} and an absorption coefficient of 0.1 cm^{-1} at 770 nm. The phantoms were imaged in a transmission geometry with the time-resolved system described in chapter 3.

Fluorescence was excited at 770 nm with a Ti:sapphire laser and detected with an 800 nm long pass filter coupled to an intensified CCD camera (PicostarHR, LAVision, Gmbh; 500 ps gate width, 600 V gain, CCD integration time of 100-200 ms, 46 time gates, 150 ps steps, 2 × 2 hardware binning). Complete TD data were recorded for up to 84 sources and 84 detectors (Fig. 4.6) arranged across the dye-filled tubes in three rows, spaced at 6.8 mm, and separated by either 1.8 mm (5.6 mm tube separation) or 1.4 mm (2 and 1.4 mm tube separation). For tomography, detectors were selected as points on the phantom surface directly above each source.

In Fig. 4.7 we display a single x–z slice from the CW, DTD, and ATD fluorescence yield reconstructions (thresholded at 90% of maximum) as well as normalized line profiles for each case.

Both the DTD and the CW approaches can resolve tubes with center-to-center separations of 5.6 mm but fail to resolve the 2 or 1.4 mm cases for any value of the regularization parameter λ. Both the 2 and 1.4 mm separations are well below the width of the typical point spread function (PSF) ($\approx$ 4–5 mm) for a 2 cm thick diffusive slab (Pogue et al., 2006) (the intrinsic PSF is also roughly represented by the reconstructed distributions, η_n for CW in Fig. 4.7). The shift of the CW reconstructions toward the long-lifetime tube for the 2 and 1.4 mm cases could be attributed to the slightly larger net yield of the 1.27 ns dye compared to the 0.88 ns dye.

Although the individual ATD distributions (η_S and η_L) are limited by the intrinsic PSF, the peaks of the distributions for each lifetime (red and green lines in Fig. 4.7) accurately represent the true tube locations for all three cases for a range of λ. The recovered tube separations for the ATD reconstructions were 5.5 ($\pm$ 0.88), 2.17 ($\pm$ 0.19), and 1.2 mm ($\pm$ 0.06 mm) at $\lambda = 1$. It is noteworthy that the DTD reconstructions used up to 18 time gates spanning the entire fluorescence decay and were thus computationally several orders more cumbersome than the ATD approach. The failure of the DTD approach to localize short separations can be attributed to the cross-talk between multiple lifetimes. However, note that the distributions for the DTD and CW are narrower than that for ATD, due to the better error performance of these methods.

4.7 Computational Efficiency

We showed in section 4.3 that both DTD and ATD approaches could be written either in a form that requires the inversion of raw TD data with the full TD weight matrix W or a form that requires inversion of the decay amplitudes with the matrix $\overline{W}$. Although for a particular approach, both forms produce the same solution, the computational efficiency and memory storage requirements for each form differ significantly. In the $\overline{W}$ form using decay amplitudes, the memory storage and total computation time have negligible dependence on number of time gates. This effec-

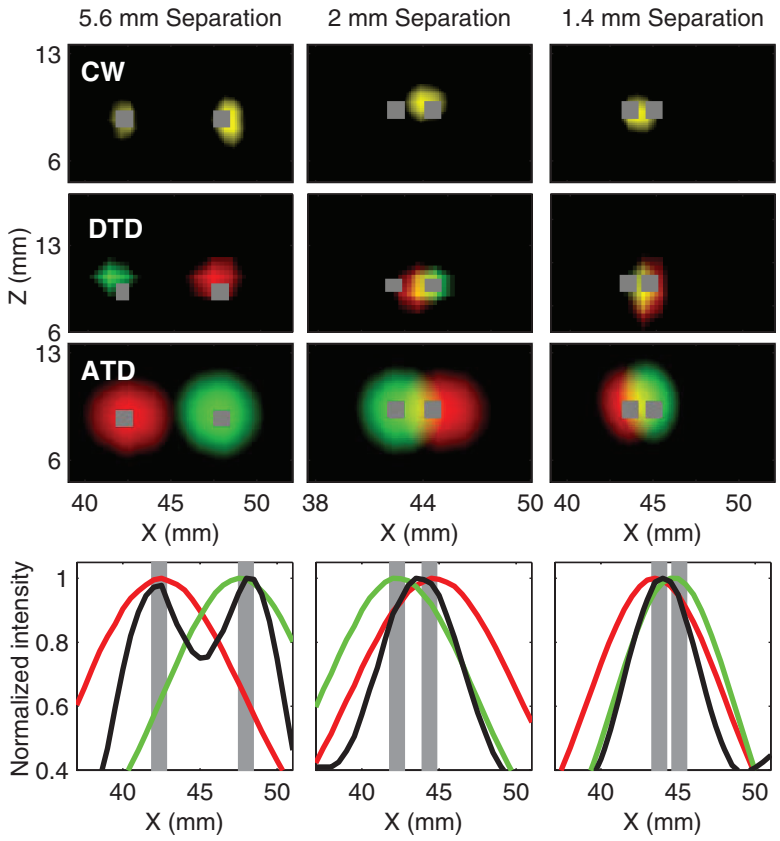

Figure 4.7: 3D fluorescence reconstructions using CW, DTD, and ATD approaches for three tube separations. Top three rows: x–z projections of the reconstructions along a central slice through the phantom. The ATD and DTD yield distributions are shown for the long (η_L) and short (η_S) lifetimes as the red and green components of an RGB image. Bottom row: corresponding normalized profiles of CW and ATD yield reconstructions along a line at a depth of 8.5 mm.

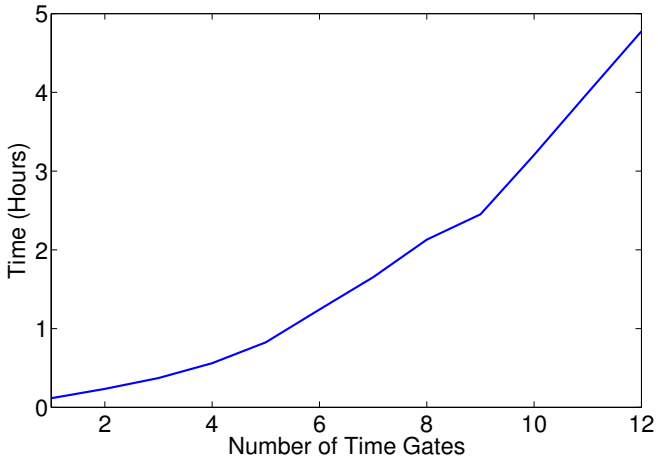

Figure 4.8: The image reconstruction time, using the full TD weight matrix W as a function of number of time gates.

tively allows the utilization of the entire decay portion of the measured TPSF for reconstruction. In contrast, for inversion of the raw TD data using the W form, the memory storage requirement shows a linear increase with number of time gates and the total computation time shows a greater than linear dependence with number of time gates.

We compare memory storage and total computation time for both forms for the simulation parameters of section 4.5. For the $\overline{W}$ form operating on decay amplitudes, the main storage term is the $\overline{W}$ matrix which occupies 5.30 GB of memory and does not depend on number of time gates. The total computation time for 50 times gates is 7.4 minutes (<1 second for multiexponential fit, 7.4 minutes for inversion of $\overline{W}$). For the W form operating on the raw TD data, the total memory storage scales as 2.65 GB $\cdot$ (number of time gates). The computation time for up to 12 times gates is shown in Fig. 4.8. It can be seen even on machines with large amounts of memory, the total computation time quickly becomes impractical even for moderate number of time gates.

4.8 Asymptotic Condition

It is known that propagation through a turbid medium distorts the fluorescence TPSF such that its decay rate is generally not the same as the intrinsic fluorescence decay rate (Patterson & Pogue, 1994). Previous studies have explored the relationship between recovered fluorescence lifetime and medium related parameters such as volume and bulk optical properties. (Patterson & Pogue, 1994) found that distortion of the recovered fluorescence lifetime is increased due to short lifetimes, thicker mediums (large source detector separations) and strong bulk scattering and weak bulk absorption. (Vishwanath & Mycek, 2004) showed a tabulation of recovered fluorescence lifetimes for different medium optical properties and thickness. (Kumar et al., 2005, 2006) established a condition for the recovery of the intrinsic lifetime for a given diffuse medium. In this section, we show that when this condition is met, an additional time constant can be derived which determines the portion of the TPSF decay that follows the intrinsic lifetime decay rate. This time constant can serve as a general approach for recovering fluorescence lifetimes from turbid tissue.

In this section, we look to determine conditions for when the fluorescence lifetime can be determined asymptotically from the fluorescence TPSF. We call the portion of the fluorescence TPSF when the time constant of the decay matches the intrinsic fluorescence lifetime the asymptotic regime. The ability to recover lifetimes asymptotically allows for a spatial-temporal separation of the fluorescence forward problem which we have applied in the asymptotic approach for reconstruction. It also allows the determination of lifetime distributions that are present deep inside tissue for lifetime-based sensing methods without having to perform forward modeling and tomographic reconstruction.

From (Patterson et al., 1989), we know that in an homogeneous semi-infinite medium, the asymptotic diffuse time scale is equal to the absorption time scale, $\tau_{abs} = 1/(v\mu_a)$. In a finite sized imaging medium, the decay time constant of the diffuse temporal response is τ_D. It has been

shown the effect of boundaries reduces the decay time constant ($\tau_D < \tau_{abs}$) (Haselgrove et al., 1992). (Kumar et al., 2005, 2006) showed the condition for asymptotic recovery of lifetimes is that the intrinsic fluorescence lifetime is longer than time constant for the temporal diffuse response ($\tau_n > \tau_D$). Given a particular lifetime τ_n can be measured asymptotically, an important time constant to consider is the time point past the peak of the fluorescence TPSF that marks the beginning of the asymptotic regime. This is also the time at which the background diffusive weight function becomes a constant. We will show that this time is characterized by a new time constant that depends on both τ_D and τ_n.

Recall that for lifetime, τ_n, we can write the TD weight function in terms of the a background weight function W_B and an exponential decay function (Kumar et al., 2006).

$$W^B(r_s, r_d, r, t) = \int_0^{t'} dt'' G^x(r_s, r, t' - t'') G^m(r, r_d, t'') \tag{4.15}$$

$$W(r_s, r_d, r, t) = \int_0^{t} dt' W^B(r_s, r_d, r, t') exp\left(-(t - t')/\tau_n\right) \tag{4.16}$$

If we assume the change in optical properties between the excitation and emission wavelengths is negligible, then both G^x and G^m decay as $\sim exp\,(-t/\tau_D)$. Hence, W^B decays as $\sim t \cdot exp\,(-t/\tau_D)$.

We rewrite (4.16) based on these time constants,

$$W(r_s, r_d, r, t) = exp\,(-t/\tau_n) \int_0^{t} dt' \alpha t' \cdot exp\left(-t'/\tau_D\right) exp\left(t'/\tau_n\right) \tag{4.17}$$

We define a new time constant τ_{BD} for the decay of the background diffuse term inside the integral:

$$\tau_{BD} = (1/\tau_n - 1/\tau_D)^{-1} \tag{4.18}$$

This time constant determines the time point at which the fluorescence TPSF enters the asymptotic regime. If τ_n is much larger than τ_D, then τ_{BD} is very close to τ_D. However, if τ_n is close to τ_D, then τ_{BD} can be very long relative to the other two time constants. In this case, even though a time point exists when the fluorescence decay term dominates, experimental conditions may make it impractical to measure data at these long time points. This can happen if the time points fall outside the repetition period of the laser or if the SNR of the time points is too low.

A simulation was performed to determine the conditions on the background absorption/scattering medium such that lifetimes can be recovered asymptotically. The simulation parameters were as follows. A 2 cm $\times$ 2 cm $\times$ 2 cm diffuse medium with background optical properties for μ_a in the range of 0 to 0.15 cm^{-1} and μ_s' in the range 10 to 50 cm^{-1} was used. A 2 mm^3 fluorescent inclusion was placed at the center of the medium. The lifetime for the inclusion was set to $\tau = 0.4$ ns. The excitation source was placed directly below the inclusion and the TPSF measurement was taken directly above the inclusion. First we determined the diffuse time constant, τ_D as a function of background optical properties. Fig. 4.9(a) is a plot with contour lines of constant τ_D. As expected, τ_D increases for decreasing bulk absorption and increasing bulk scattering. Next we fit for the fluorescence lifetimes starting from t_{fit} which is chosen to be the time point past the peak of the fluorescence TPSF corresponding to 0.2% of the peak value. Fig. 4.9(b) shows a plot with contour lines of constant recovered lifetimes. It can be seen that lifetimes are generally recovered for bulk optical properties corresponding to τ_D lower than τ. By choosing a set time point $t_{\hat{fit}}$ to begin the fitting procedure, we mimic the actual experimental situation where due to SNR and other constraints we cannot choose time points arbitrarily past the peak. However, since $t_{\hat{fit}}$ may not meet the condition $t_{\hat{fit}} \gg \tau_{BD}$, it is possible that the true lifetime is not recovered even though the asymptotic condition holds ($\tau > \tau_D$).

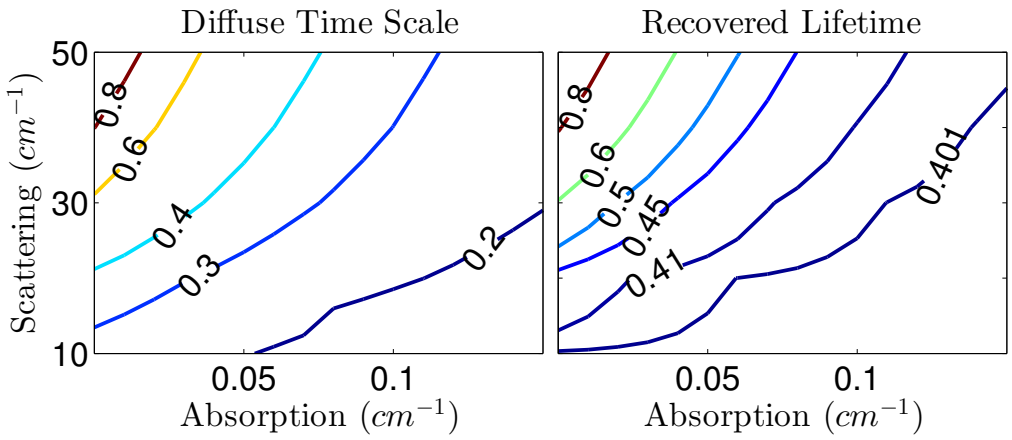

Figure 4.9: Contour plots (units of ns) for the diffuse time scale, τ_D (left) and recovered lifetime, τ (right) for varying bulk optical properties of the medium μ_a and μ_s'.

4.9 Timing Jitter Sensitivity

In this section, we show how accurate determination of the start of the asymptotic regime has practical importance for reconstruction of experimental data. We saw in chapter 3 that timing jitter is an experimental impediment for time-domain systems. We will show below that image reconstruction in the asymptotic regime is less sensitive to timing jitter than reconstruction from time points before the asymptotic regime. If we assume the time origin for the measurement weight matrix is t_0 and the time origin for the model weight matrix has undergone a shift of δ_{jitter} (this can occur because the reference IRF and IRF of the TD data are taken at different times and the reference IRF has undergone a shift relative to the IRF of the TD data) then the measurement and model weight matrices within the asymptotic regime are:

$$\text{Measurement: } Wexp(-(t-t_0)/\tau) \tag{4.19}$$

$$\text{Model: } Wexp(-(t-t_0-\delta_{jitter})/\tau) \tag{4.20}$$

and outside the asymptotic regime are:

$$\text{Measurement: } W(t - t_0)exp(-(t - t_0)/\tau) \tag{4.21}$$

$$\text{Model: } W(t - t_0 - \delta_{jitter})exp(-(t - t_0 - \delta_{jitter})/\tau) \tag{4.22}$$

It can be seen that timing jitter only introduces a scaling factor ($exp(\delta_{jitter})$) for the reconstruction problem in the asymptotic regime without changing the spatial distribution of the reconstruction while timing jitter will change the spatial distribution of the reconstruction for TD data not in the asymptotic regime.

5

Hybrid Reconstruction Techniques for Tomographic Lifetime Imaging

It has been shown that reconstruction using direct time points from early TD data can provide higher spatial resolution than CW and late TD data (Niedre et al., 2008), due to the minimal scattering experienced by early arriving photons. However, the early TD data offers poor lifetime sensitivity, resulting in significant lifetime cross-talk and inaccurate localization of closely sepa-

rated fluorophores with distinct lifetimes (Kumar et al., 2006). In chapter 4, we showed that an asymptotic TD (ATD) approach based on a multi-exponential analysis of the decay portion of TD data provides low cross-talk and superior tomographic localization of multiple fluorophores with distinct lifetimes. The ATD approach relies on the asymptotic factorization of the TD fluorescence into diffusive and pure exponential fluorescence decay components (Kumar et al., 2006).

In this chapter we present two novel, hybrid TD approaches that combine the resolving power of early photon tomography with the lifetime multiplexing capability of the asymptotic approach thereby achieving high resolution tomographic lifetime multiplexing. Experimental results from phantom measurements demonstrate that the hybrid method is able to separate targets as close as 1.4 mm apart with minimal cross-talk and with distributions narrower than that obtained from any individual method.

5.1 Early Photon Tomography

Early photon tomography (EPT) is a technique that has been explored as a way to improve the spatial resolution of diffuse optical tomography (DOT) and fluorescence DOT. In EPT, the effect of tissue scattering is reduced by considering early arriving photons which preferentially follow straighter paths from source to detector. This reduction in scattering can lead to a better conditioned inverse problem (Leblond et al., 2009). However, the loss in SNR associated with neglecting diffusive photons has also been found to result in lower detection sensitivity as compared to CW imaging (Valim et al., 2010). Despite this potential drawback, EPT has been applied successfully for in vivo monitoring of tumor growth in a lung carcinoma model where it was found to offer better resolution than conventional CW imaging (Niedre et al., 2008).

To demonstrate resolution improvements offered by early photons, we plot for a single source and detector, the time dependent weight function (or sensitivity) for early time points on the flu-

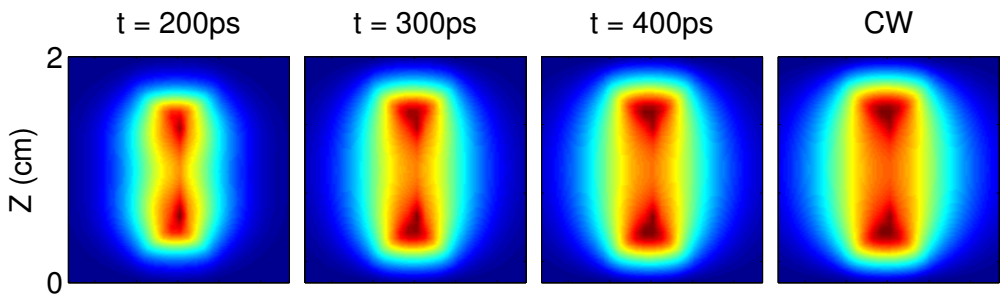

Figure 5.1: Spatial extent of the TD sensitivity function for t = 200 ps, 300 ps, and 400 ps and of the CW sensitivity function.

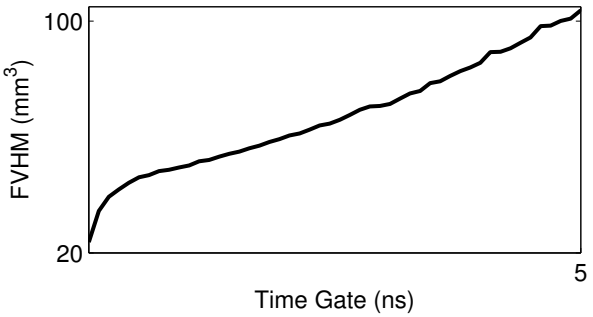

Figure 5.2: Resolution as a function of time gate, measured by the full volume at half maximum of the reconstruction of a point inclusion at the center of a 2 cm turbid medium.

orescence TPSF in Fig. 5.1. As a comparison, the CW weight function is also plotted in the same figure. It is clear that the spatial spread of the sensitivity function is significantly reduced for early time gates compared to the CW sensitivity function indicating the less diffusive nature of early photons. In Fig. 5.2, we quantify the resolution by plotting the recovered full volume at half maximum (FVHM) for a fluorescent inclusion at the center of a diffuse medium for each time gate. For these simulations, realistic shot noise is added to the data. It can be seen despite the lower SNR of early time gates, the FVHM increases with later time gates.

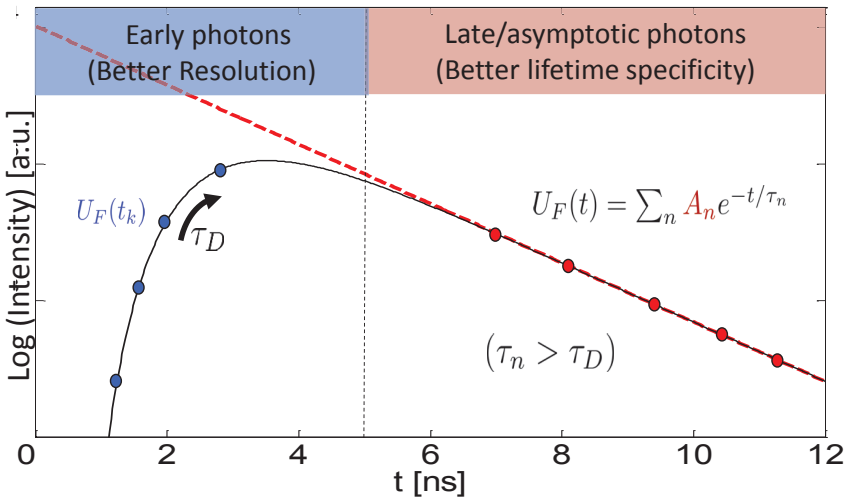

Figure 5.3: The two parts of the fluorescence temporal point spread function. Early photons are shown in blue and consists of points with highest SNR and which can provide highest resolution. Late photons are shown in red and can be modeled as a weighted exponential sum.

5.2 Hybrid Reconstruction Techniques

In this section, we present two novel techniques to combine early and late arriving photon data. Fig. 5.3 shows representative early and late time gates on the fluorescence TPSF. For convenience, we restate the two forms for the inversion of time domain data below:

$$\begin{aligned} DTD : y = W\eta \qquad & \Longleftrightarrow \qquad \eta^{DTD} = W^T(WW^T + \lambda I)^{-1}y \text{ (for early and late time gates)} \\ ATD : y = A\overline{W}\eta \qquad & \qquad\qquad \eta^{ATD} = \overline{W}^T(\overline{W}\overline{W}^T + \lambda I)^{-1}A^{\dagger}y \text{ (for late time gates)} \end{aligned}$$

The individual components for each hybrid technique are as follows:

- For time gates corresponding to the earliest to peak portions of the TPSF, we will apply a DTD approach for reconstruction. This corresponds to the time points where the highest spatial resolution can be achieved.

- For late time gates, we will apply an ATD approach for reconstruction. We showed in chapter

4 that ATD results in significantly less cross-talk than inverting the full TD sensitivity matrix. Also, initially performing a linear fit in the time domain before applying a regularized inverse with CW weight matrices can significantly reduce the reconstruction time and memory requirements compared to a direct inversion.

5.2.1 Asymptotic Reconstruction as Spatial Prior

A unique feature of the TD weight matrix in the asymptotic regime is that the decay profile for each lifetime component is independent of voxel location. The total TD measurement data for these time points can be readily separated into temporal terms which correspond to each lifetime. On the other hand, the TD weight matrix for an early times gate consists of different temporal profiles at each voxel. For a TD measurement at these time gates, it is no longer possible to determine the temporal terms that correspond to each lifetime component without consideration for the spatial distribution of the fluorophores. For early time gates, instead of attempting to separate different fluorophores using TD measurement data, we consider separation of the fluorophores at the voxel level using reconstructions derived from late time gates. We implement this separation in the form of a spatial prior.

Spatial priors have previously been applied in optical tomography by incorporation of a low resolution estimate of the imaging object, typically from another imaging modality such as magnetic resonance imaging (MRI) and x-ray as prior information in the optical reconstruction (Li et al., 2003, 2005). For our application, an initial lower resolution, but well separated, estimate of the fluorescence yield is obtained using the ATD approach applied to late gates (η_{ATD}). Next, reconstruction on early time gates is performed with the spatial regularization matrix $L = diag(1/\eta_{ATD})$ encoded with the ATD reconstruction:

$$\eta = (W^T W + \lambda L^T L)^{-1} W^T y \tag{5.1}$$

This method will subsequently be referred as the Hybrid TD Asymptotic Prior (HTD-AP) method.

5.2.2 Combined Matrix with Multiple Regularization Parameters

In order to incorporate the advantages of early gates with the ATD approach into a single self consistent inverse problem, we consider inverting a "hybrid" data set consisting of early time gates and decay amplitudes, $\mathcal{Y} = [y(t_1), y(t_2), ..., a_1, a_2, ...]^T$, using a hybrid weight matrix, $\mathcal{W}$ consisting of the DTD and ATD weight functions (W_n and $\overline{W}_n$). The hybrid TD (HTD) forward problem takes the form $\mathcal{Y} = \mathcal{W}\eta$ or explicitly,

$$\begin{bmatrix} y(t_1) \\ y(t_2) \\ \vdots \\ a_1 \\ \vdots \\ a_N \end{bmatrix} = \begin{bmatrix} W_1(t_1) & \dots & W_N(t_1) \\ W_1(t_2) & \dots & W_N(t_2) \\ \vdots & \dots & \vdots \\ \overline{W}_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \overline{W}_N \end{bmatrix} \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_N \end{bmatrix}. \tag{5.2}$$

(5.2) can be inverted using Tikhonov regularization:

$$\eta = \mathcal{W}^{\mathcal{T}}(\mathcal{W}\mathcal{W}^{\mathcal{T}} + \lambda\mathcal{I})^{-1}\mathcal{Y} \tag{5.3}$$

It is advantageous to impose a positivity constraint (see also (Bertero & Boccacci, 1985)) on the Tikhonov cost functional, so that the inversion of the HTD approach is expressed as the minimization problem:

$$\eta^{HTD} = \arg\min_{\eta \geq 0} \|\mathcal{Y} - \mathcal{W}\eta\|^2_{C^{-1}} + \lambda\|\eta\|^2, \tag{5.4}$$

where C is a measurement covariance matrix that incorporates the distinct noise characteristics of the direct time points y and the amplitudes a within $\mathcal{Y}$. The positivity constraint reduces cross-talk with the HTD approach by preventing negative yield values from spuriously reducing the ATD term of the HTD cost functional in (5.4). This method will subsequently be referred as the Hybrid TD Combined Matrix (HTD-CM) method.

5.3 Simulations - Rectangular Slab

We investigated the performance of hybrid algorithms for point inclusions and complex and overlapping shapes in a rectangular slab. Early time gate, ATD and HTD reconstructions are obtained for each problem. For the most general problem of two point inclusions of different lifetimes we show results for both hybrid methods (HTD-AP and HTD-CM). For other simulation problems, we present the results for only one of the hybrid methods. We quantified the theoretical results of section 5.2 using Monte Carlo simulations, performed using tMCimg (Boas et al., 2002). The simulation medium for the point inclusions and complex shapes was a 2 cm $\times$ 2 cm $\times$ 2 cm diffuse box with bulk optical properties $\mu_a = 0.1\ \text{cm}^{-1}$ and $\mu_s' = 10\ \text{cm}^{-1}$. 49 sources and 49 detectors were located at the z = 0 cm and z = 2 cm planes respectively. Sources and detectors covered the medium surface in a uniform grid configuration with distance between adjacent sources/detectors to be 2 mm. Minimization of the HTD cost functional with positivity constraint was implemented using L-BFGS-B (Byrd et al., 1995) in MATLAB (The Mathworks, Inc.). In all simulations, the regularization parameter was chosen corresponding to the least reconstruction error ($E = \|\eta_{recon} - \eta_{true}\|$).

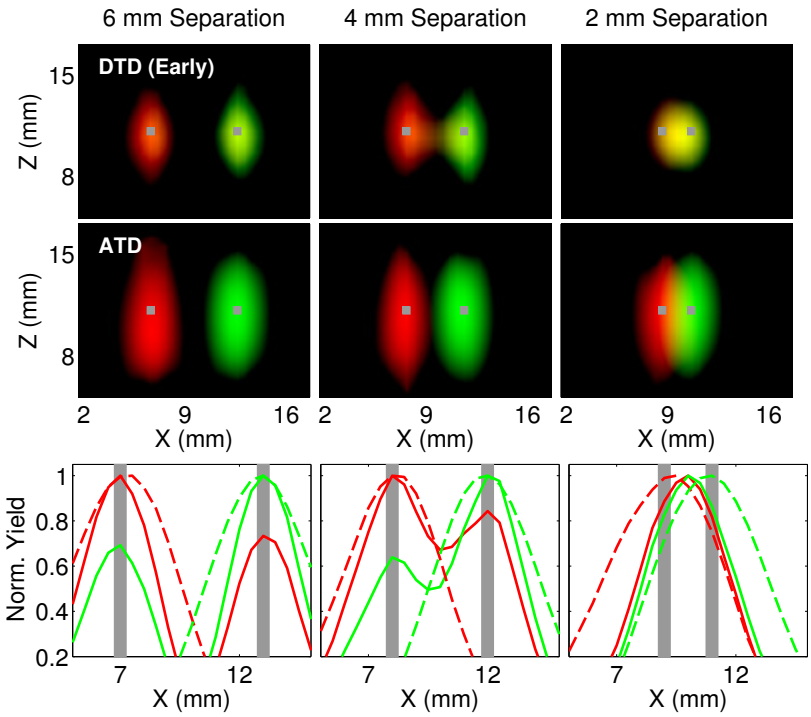

Figure 5.4: Comparison of reconstructions obtained from applying ATD to 12 late time gates and DTD to 4 early late time gates. Red and green corresponds to the yield distributions for 0.87 and 1.27 ns, respectively. The true locations of the inclusions are shown in gray. In the first two rows, the X–Z plots are generated by assigning the recovered yields to the red (0.87 ns) and green (1.27 ns) components of the RGB colormap. Each distribution is thresholded at 50% of its maximum. The bottom row shows line plots for ATD (dashed line) and DTD (solid line) along the x-axis at the depth of the inclusion.

5.3.1 Two point inclusions

Two fluorescent inclusions of lifetimes $\tau_1 = 0.87$ ns and $\tau_2 = 1.27$ ns were embedded at the z = 1 cm depth and separated by 2 mm, 4 mm and 6 mm. We first evaluate the individual performance of DTD approach for early time gates and the ATD approach for late time gates. Figure 5.4 shows the ATD approach is able to separate the two inclusions with minimal cross-talk. Figure 5.4 also shows reconstructions using the DTD approach with four early time gates. While the yield distributions are smaller than the ATD distributions, the cross-talk is significantly higher than ATD for all separations leading to inaccurate localization.

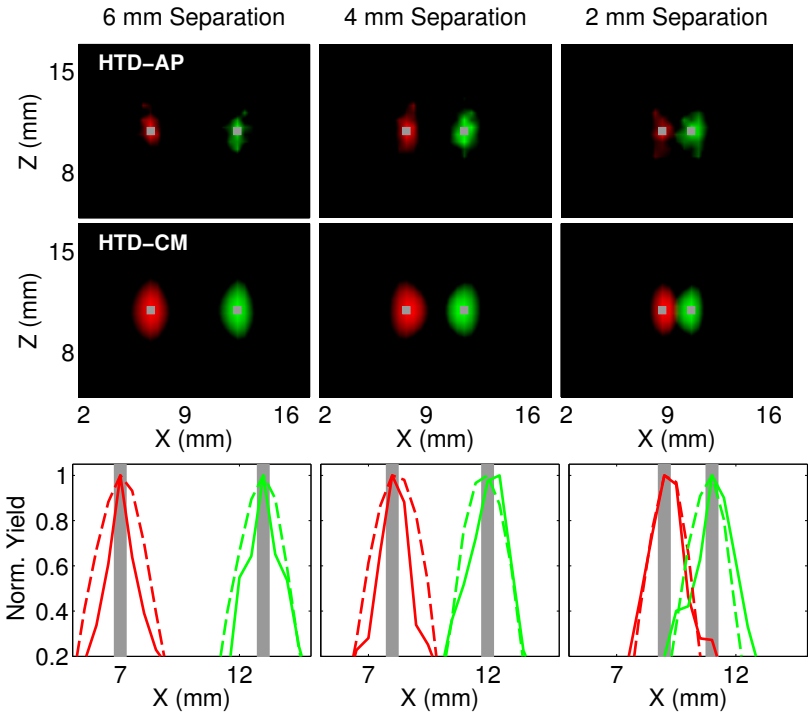

Figure 5.5: Reconstruction obtained from both HTD methods. Red and green correspond to the yield distributions for 0.87 ns and 1.27 ns respectively. Each distribution is thresholded to 50% of its maximum. The bottom row shows line plots for HTD-AP (solid line) and HTD-CM (dashed line) along the x-axis at the depth of the inclusions.

We next applied the hybrid approaches for reconstruction. Figure 5.5 shows the reconstructions using the HTD-CM approach with four early time gates and the decay amplitudes, which provides high resolution and accurate localization for all separations. We note that when the positivity constraint (5.4) was not imposed, the cross-talk of the HTD-CM method was higher, but still lower than for DTD alone. Figure. 5.5 also shows the HTD-AP reconstruction as applied to four early time gates using the asymptotic reconstruction as a spatial prior. It can be seen that separation and localization of the inclusions could also be achieved down to 2 mm.

5.3.2 Four point inclusions

Simulations with two point inclusions of differing lifetimes showed that objects separable by the ATD method could be made better resolved using information from early time gates. However, in the case of multiple closely spaced objects of different lifetimes, the HTD method has the potential to resolve objects not resolvable by using either the lifetime contrast from ATD or the resolving power of early time gates alone.

Four point inclusions in a line are embedded at the center of the diffuse medium. The two left-most inclusions have a lifetime of ($\tau = 0.87$ ns) while the two right-most inclusions have a lifetime of ($\tau = 1.27$ ns). The inclusions are spaced such that the inner inclusions are separated by 2 mm while the outer inclusions on both sides are separated by 4 mm. Even though ATD can resolve each pair of inclusions with different lifetimes, it cannot resolve inclusions within each pair with the same lifetime (Fig. 5.6). Early time gates cannot differentiate the lifetimes of the inclusions and cannot resolve the two inner inclusions. However, it can be seen that the HTD approach can correctly localize and resolve all four inclusions.

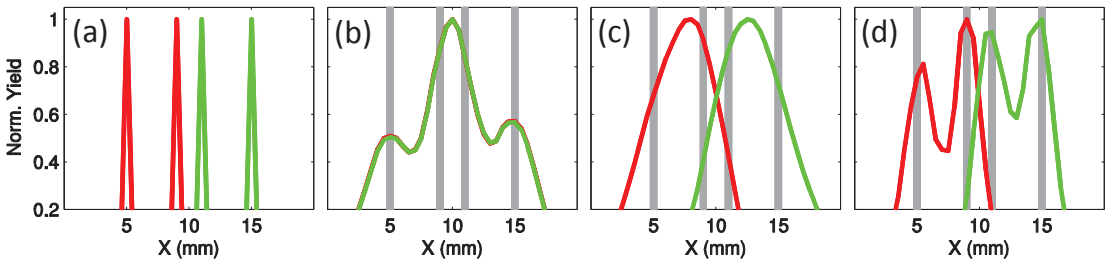

Figure 5.6: Resolving closely spaced objects of differing lifetimes. (a) Four fluorescent inclusions placed in a line with two inner inclusions separated by 2 mm and two outer inclusions on both sides separated by 4 mm. $\tau = 0.87$ ns is shown in red while $\tau = 1.27$ ns is shown in green (b) Early time gate reconstruction (c) ATD reconstruction (d) HTD reconstruction.

5.3.3 Complex Shapes

Although point inclusions are a simple model to evaluate the performance of the hybrid reconstruction algorithms, simulations of larger objects with complex shapes would provide a more realistic representation of how the algorithm would perform in small animal imaging applications such as the imaging of deep-seated tumors. We chose three test cases that can illustrate the ability of the algorithm to separate closely spaced fluorescent objects with varying degrees of overlap. The test cases were: adjacent semi-circles, overlapping ellipses and concentric circles in a slab model. The first column of Fig. 5.7 shows x-y projections of the true objects. For the semi-circles, the left half had $\tau_1 = 0.87$ ns while the right half had $\tau_2 = 1.27$ ns. The diameter of the semi-circles was 1 cm. For the overlapping ellipses, the ellipses with the major axis aligned with the x-axis and y-axis had lifetimes of $\tau_1 = 0.87$ ns and $\tau_2 = 1.27$ ns respectively. Both ellipses had major axis length of 1 cm and minor axis length of 4 mm. For the concentric circles, the larger circle had a diameter of 1 cm and $\tau_1 = 0.87$ ns while the smaller circle had a diameter of 5 mm and $\tau_2 = 1.27$ ns. For all three test cases, the objects were located at the center of the medium. The objects were uniform along the z-direction and had thicknesses of 5 mm.

It can be seen that the early gate reconstructions showed a strong cross-talk component for all three cases. This was particularly noticeable for the semi-circles where a strong overlap between

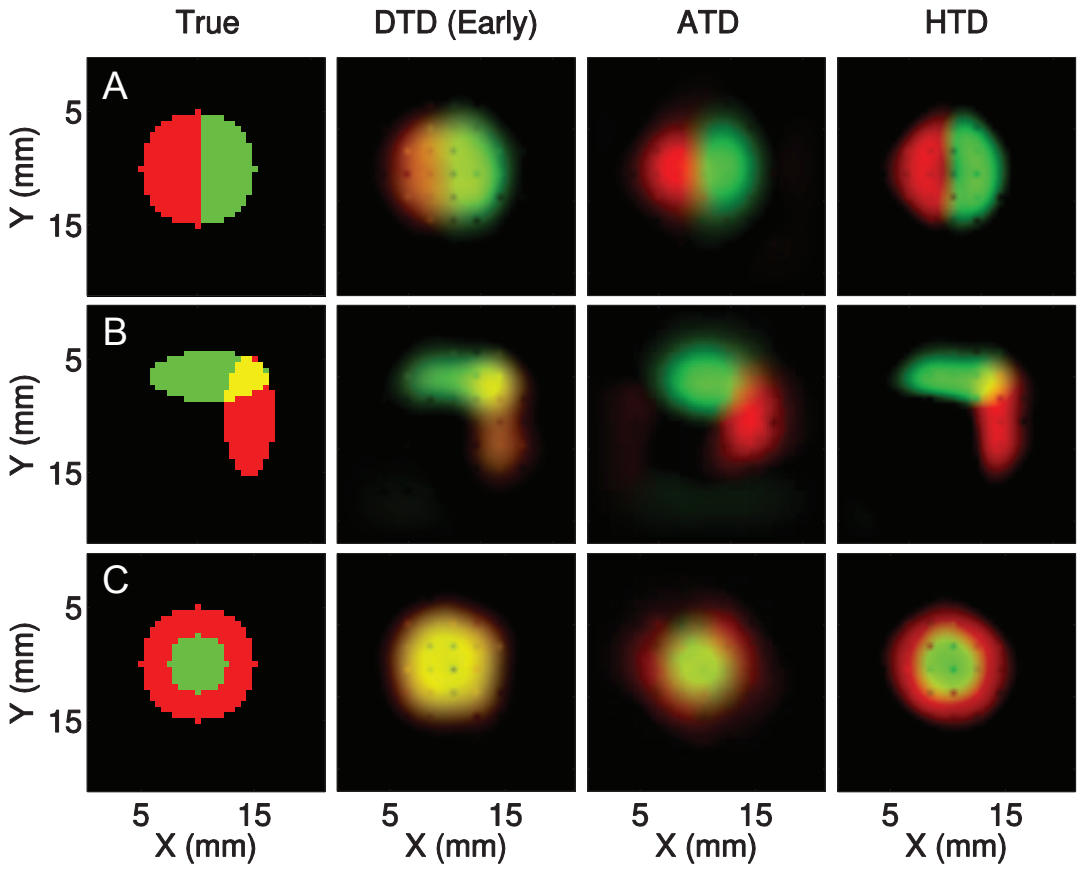

Figure 5.7: Reconstruction of complex and overlapping shapes. The true objects for three test cases, A - adjacent semi-circles, B - overlapping ellipses and C - concentric circles are shown in column 1. Red and green correspond to the yield distributions for 0.87 ns and 1.27 ns respectively. In all test cases, objects with a thickness of 5 mm are embedded in the center of a 2 cm thick slab. Reconstructions using early time gates (column 2), ATD method (column 3) and HTD method (column 4) are shown for all test cases.

the two halves was present and for the concentric circles where the presence of a surrounding circle is no longer apparent. The ATD reconstruction showed improved separation for all three test cases compared to early time gates even though the object shapes were not well preserved. This distortion of the shape made it so that the overlap between the two lifetime components was not recovered for the overlapping ellipses. While for the concentric circles, the ATD reconstruction could recover the relative size difference of the circles, the hole shape of the surrounding circle was not well captured. Finally, HTD reconstructions showed the sharpest images with the least amount of cross-talk compared to DTD of early time gates and ATD alone for all three test cases. The shapes for the semi-circles and overlapping ellipses were well recovered using the HTD approach. Also for the concentric circles test case, the inner circle could be easily be distinguished from the surrounding circle.

5.4 Simulations - Mouse Atlas

We next explore the feasibility of using hybrid reconstruction techniques for lifetime tomography in mice by considering two potential applications: anatomical imaging using organ targeting dyes and cardiac imaging with an activatable probe.

Anatomical imaging using organ targeting dyes with lifetime contrast would allow for the validation of 3D reconstructions using fluorescence sources inside a living mouse with complex shapes and known locations. Also, the ability to optically image organs would provide an opportunity for all-optical methods for fluorescence tomography with anatomical priors. Currently anatomical landmarks such as the skeleton are obtained using other imaging modalities such as CT or MRI (Grimm et al., 2005). An all-optical method for combined functional and anatomical imaging would eliminate the need of performing co-registration between different imaging modalities. In this study, we consider lifetime multiplexing with two organs in close proximity to

each other, the skeleton and the kidneys. In vivo reconstructions for dyes targeting the kidneys and skeleton are presented in chapter 8.

A major problem in imaging targeted fluorescent probes is background accumulation of flurophores, especially in the liver. This problem can be addressed by using activatable probes whose lifetime shifts upon activation. We have shown previously the ability to detect protease activation in infarcted myocardisum with an activatable near-infrared probe (PGC-800) (Goergen et al., 2012). Planar fluorescence imaging showed that after infarction the lifetime in the thoracic region (0.53 ns) shifted relative to the lifetime in the liver (0.67 ns). Tomographic reconstruction based on the amplitude of the thoracic signal would allow for quantitative estimates of the in vivo concentration of the activated probe.

Simulations were performed in a publicly available digital mouse atlas (Dogdas et al., 2007) generated using CT and cryosection data from a nude mouse. The atlas is segmented into 21 anatomical regions with each voxel having a corresponding region label. For our study, the original 0.1 mm voxels was binned down to 1 mm resulting in a $38 \times 99 \times 21$ volume. Voxels corresponding to the skeleton and kidneys were assigned lifetimes of $\tau_1 = 0.5$ ns and $\tau_2 = 0.85$ ns respectively while the voxels corresponding to the heart and liver were assigned lifetimes of $\tau_1 = 0.53$ ns and $\tau_2 = 0.67$ ns respectively. The fluorescence yield for all objects were set to be the same value. Simulated TD measurements were generated for 87 source/detector pairs across the mouse. Measurements were convolved with an experimentally derived instrument response function assuming an image intensifier gate width of 500 ps. 2% shot noise was added to all data prior to reconstruction. Both the true objects and reconstructions are displayed in Fig. 5.8 and 5.9 as maximum intensity projections. For cardiac imaging, the true skeleton was included as the blue component in all images to show the relative locations of the true and reconstructed organs. For anatomical imaging, reconstructions are shown for the horizontal plane while for cardiac imaging, reconstruction are shown for the sagittal plane. CW reconstructions were also included

in the comparison to show the potential benefits of using TD over CW data for reconstruction.

From Fig. 5.8, it can be seen that when imaging the skeleton and kidneys using CW, the relative magnitude of the objects could not be recovered accurately. This is due to cross-talk between the lifetime components and makes CW reconstructed yields in a medium with multiple lifetimes unreliable. Even though the early time gate reconstructions was able to delineate the shapes of the organs quite well, there was significant cross-talk especially within the kidneys. ATD showed excellent separation of the organs as expected although the shapes of the kidneys and the skeleton appeared more blurred relative to the early time gate reconstruction. HTD reconstructions showed complete separation of the two organs. In particular, it can be observed only the HTD reconstruction could recover the segment of the skeleton in the region between the two kidneys.

For cardiac imaging, the quality of the reconstructions (Fig. 5.9) followed a similar pattern to the skeleton-kidneys reconstructions. CW reconstructions were able to differentiate the two organs due to the difference in lifetime of the probe in both regions. However, the relative increase of the reconstructed yield in the heart relative to the liver did not match the yields of the true objects. This can be attributed to the non-uniqueness of lifetime and fluorescence yield for CW reconstructions. Early time gate reconstructions were able to provide moderate amount of separation of the heart from the liver, possibly due to separated nature of the two organs but a significant cross-talk component is still visible. The ATD reconstruction showed clear separation of the two organs although the shape of liver is not well recovered. Finally, HTD was able to separate both organs with high resolution.

5.5 Experiments

We validated the HTD approach using phantom experiments. Parallel tubes filled with two dyes (τ = 0.87 and 1.27 ns) and separated by 5.6, 2, and 1.4 mm were embedded in a scattering

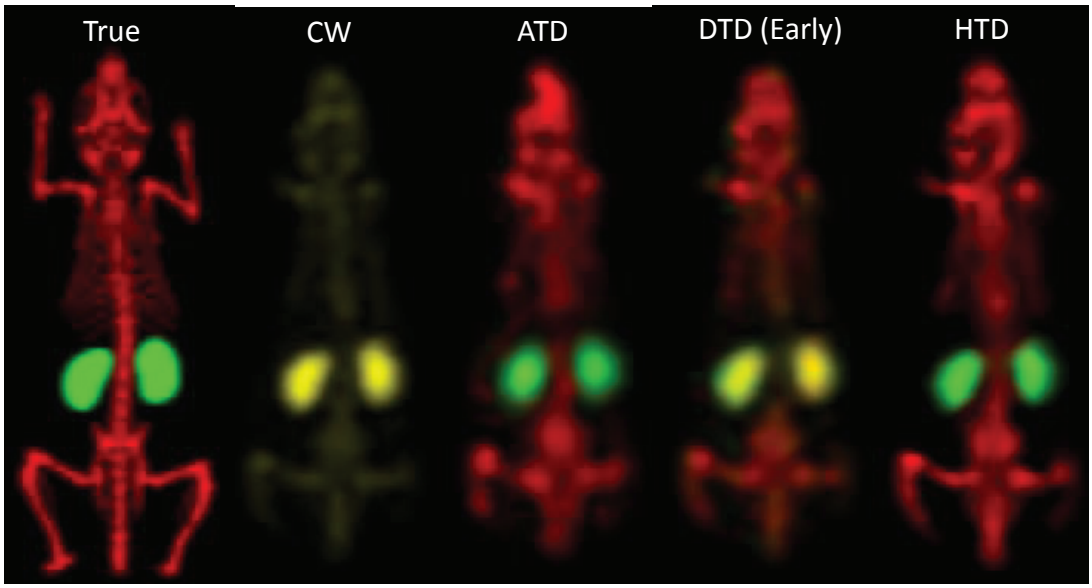

Figure 5.8: Tomographic lifetime multiplexing of organ targeted fluorophores, using the Digimouse atlas with 84 sources and 84 detectors. 3D reconstructions were performed on simulated TD data which was convolved with an experimentally determined IRF (gatewidth = 500 ps). The lifetimes of the skeleton and kidneys were $\tau = 0.5$ ns and $\tau = 0.85$ ns respectively. The true yield distribution and reconstructed yield distributions for CW, early time gates, ATD and HTD methods are shown using the are shown for lifetimes $\tau = 0.5$ ns and $\tau = 0.85$ ns as the red and green components of an RGB image.

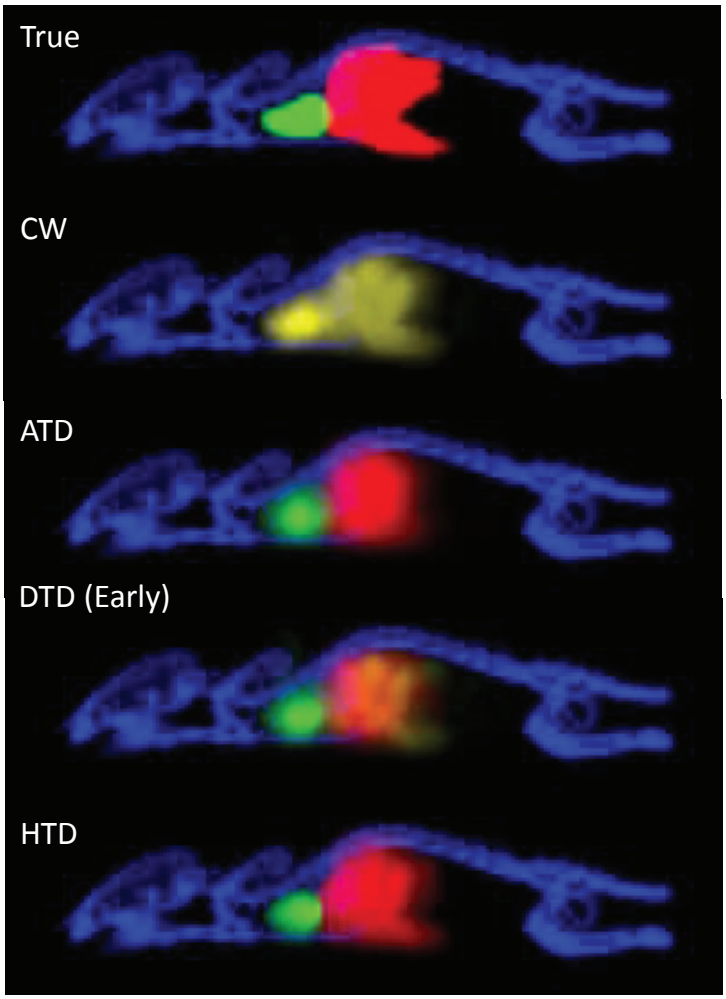

Figure 5.9: Tomographic lifetime multiplexing applied to cardiac imaging with an activatable probe, using the Digimouse atlas with 84 sources and 84 detectors. 3D reconstructions were performed on simulated TD data which was convolved with an experimentally determined IRF (gatewidth = 500 ps). The lifetimes of the heart and liver were $\tau = 0.53$ ns and $\tau = 0.67$ ns respectively. The true yield distribution and reconstructed yield distributions for CW, early time gates, ATD and HTD methods are shown using the are shown for lifetimes $\tau = 0.53$ ns and $\tau = 0.67$ ns as the red and green components of an RGB image.

medium consisting of Intralipid and nigrosin ($\mu_a = 0.1\ \mathrm{cm}^{-1}$ and $\mu_s' = 10\ \mathrm{cm}^{-1}$). Measurements were performed with a previously described TD fluorescence tomography system (Kumar et al., 2008b), consisting of a Ti-Sapphire laser for excitation and time-gated intensified CCD camera for detection. Full tomographic measurements were acquired for up to 84 sources and 84 detectors, and 46 time gates, with a gate width of 500 ps, CCD integration time of 100–200 ms, and step size of 150 ps.

The results shown in Fig. 5.10 indicate that the ATD method is able to correctly localize the tubes for all separations, while DTD and CW cannot separate tubes closer than 5.6 mm. However, the HTD method applied to three early time gates and decay amplitudes is able to accurately localize the inclusions, with significantly narrower yield distributions compared to the ATD method. We note that the same regularization was applied to ATD alone as to the ATD component of the hybrid method, indicating that the improvement in resolution is due to the incorporation of early time gates in the reconstruction.

In summary, we have presented a novel approach for TD fluorescence tomography that combines the use of early- and late-arriving photons, enabling high-resolution lifetime tomography in turbid media. We have verified the approach in simulations involving simple point inclusions and more complex shapes and organ labeling in a digital mouse atlas. We have experimentally demonstrated the improved resolution and cross talk performance of the hybrid method in separating closely located targets. The spatial resolution of the HTD approach can be further improved using faster detectors that are capable of detecting photons arriving earlier than 100 ps. Future work will be focused on in vivo applications of this technology in animal models of disease.

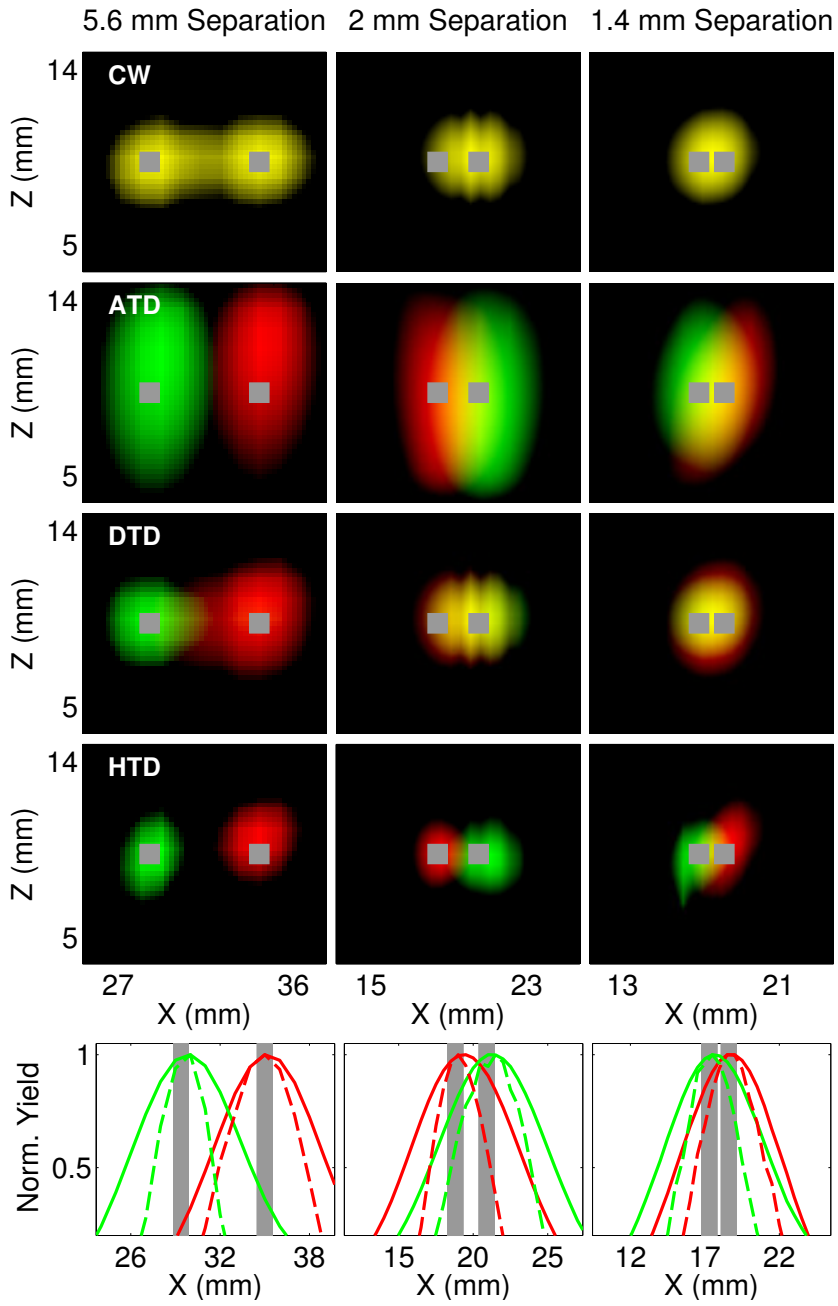

Figure 5.10: Experimental reconstruction of fluorophores with lifetime contrast in a dish phantom. X–Z and line plots for CW, ATD, DTD (three early and three late gates), and HTD combining ATD and DTD (three early gates) are shown. Red and green correspond to the yield distributions for 0.87 and 1.27 ns, respectively. Each distribution is thresholded at 70% of its maximum. The bottom row shows line plots for HTD (dashed line) and ATD (solid line) along the x-axis at the depth of the inclusion.

6

Comparison of Tomographic Spectral and Lifetime Multiplexing

Mutltispectral methods have commonly been used in microscopy to separate multiple fluorophores in a tissue sample using their unique spectral signatures (Zimmermann et al., 2003). These methods have also been successfully applied to distinguish fluorophores of interest from tissue autofluorescence (Mansfield et al., 2005; Levenson et al., 2008). In microscopy, the unmixing can be

performed pixel by pixel with simple linear fits to predetermined spectral basis functions. However, in the case of thick samples, the measurement at the surface of the sample is distorted by the wavelength dependent effects of light propagation (Chaudhari et al., 2009; Leblond et al., 2011). Quantitative separation of fluorophores embedded deep in tissue using multispectral imaging requires accurate modeling of absorption and scattering at multiple wavelengths and computationally inverting the effects of propagation through the medium.

In this chapter, we show that the multispectral fluorescence tomography inverse problem can be written in two forms analogous to TD lifetime multiplexing. We also demonstrate that the asymptotic TD (ATD) approach is mathematically distinct from multispectral tomography in a manner that offers unique advantages over spectral methods for separation of fluorophores in turbid media. We use near-infrared fluorescent proteins as a model to compare the performance of lifetime and spectral methods.

6.1 Mathematical Form of Multispectral Fluorescence Tomography

From chapter 2, we know the propagation of photons in fluorescence tomography can be modeled by coupled radiative transport equations at the excitation and emission wavelengths. Under the assumption of at most a single excitation/emission event per photon, the surface fluence can be written as a linear function of the fluorescence yield. In the case of multispectral fluorescence tomography (MSFT), the measurements are taken at multiple wavelengths and fluorescence yield is a function of wavelength resulting in the following forward equation:

$$U(\mathbf{r}_s, \mathbf{r}_d, \lambda_x, \lambda_m) = \int_\Omega G^x(\mathbf{r}_s, \mathbf{r}, \lambda^x) G^m(\mathbf{r}, \mathbf{r}_d, \lambda^m) \eta(\mathbf{r}, \lambda^x, \lambda^m) d^3 r, \tag{6.1}$$

where G^x is the source and G^m is the detector Green's functions evaluated for optical properties at the excitation and emission wavelengths, λ_x and λ_m respectively and $\eta(r)$ is the fluorescence

yield. In excitation-resolved fluorescence tomography, measurements are taken at different λ_x and collected for a single λ_m while in emission-resolved fluorescence tomography λ_x is held constant while measurements are taken at different emission wavelengths. For both measurement modes, η can be factored into a spectral term $\epsilon(\lambda)$ containing the quantum yield, extinction coefficient and relative emission strength and a spatial term containing the unknown concentration $C(r)$. For simplicity, we consider the emission-resolved case below where only $\lambda_m = \lambda$ is varying. (6.1) can be rewritten as:

$$U(\mathbf{r}_s, \mathbf{r}_d, \lambda) = \int_\Omega G^x(\mathbf{r}_s, \mathbf{r}, \lambda^x) G^m(\mathbf{r}, \mathbf{r}_d, \lambda) \epsilon(\lambda) C(\mathbf{r}) d^\beta r, \tag{6.2}$$

We now consider the problem of fluorophore multiplexing using multispectral data where the aim is to separate N fluorophores with concentrations $C_n(r)$ and spectral terms $\epsilon_n(\lambda)$ using data at M source/detector pairs and L wavelengths. The fluorescence yield in (6.1) can first be replaced by a spectrally weighted sum of the concentrations:

$$\eta(r, \lambda) = \sum_{n=1}^{N} \epsilon_n(\lambda) C_n(r) \tag{6.3}$$

We discretize the medium into V voxels so that (6.2) and (6.3) can be written as a linear matrix equation:

$$y = Wc \tag{6.4}$$

where W is a $(ML \times VN)$ multispectral weight matrix containing both information about the wavelength dependent absorption/scattering properties of the medium through the Green's functions and the emission spectra of the fluorophores. y is a $(ML \times 1)$ vector containing the measurement at different emission wavelengths and c is a $(VN \times 1)$ vector containing the concentrations for all fluorophores.

The full weight matrix W can be also be factored into spatially dependent and spectrally depen-

dent matrices so that (6.4) can be rewritten as:

$$y = \overline{W}Ac \tag{6.5}$$

$$\overline{W} = \begin{bmatrix} \overline{W_1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \overline{W_N} \end{bmatrix}, A = \begin{bmatrix} \epsilon_1(\lambda) * I & \dots & \epsilon_N(\lambda) * I \end{bmatrix} \tag{6.6}$$

where $\overline{W}$ is a $(ML \times VL)$ matrix containing CW weight matrices at each emission wavelength, I is a (V $\times$ V) identity matrix and A is a $(VL \times VN)$ matrix containing the spectral basis function for each fluorophore.

To reconstruct for the unknown concentration C_n, two methods will be considered. In the direct approach, the full weight matrix in (6.4) is inverted using Tikhonov regularization:

$$c^{Direct} = W^T(WW^T + \alpha I)^{-1}y \tag{6.7}$$

In the indirect approach, the inversion is performed on (6.5) in two steps. First, $\overline{W}$ is inverted using Tikhonov regularization. Next, due to the well conditioned nature of A, it is inverted without regularization by multiplication with its Moore-Penrose pseudoinverse. This is equivalent to performing a least squares fit at each voxel using the spectral basis functions, $\epsilon_n(\lambda)$.

$$c^{Indirect} = A^{\dagger}\overline{W}^T(\overline{W}\overline{W}^T + \alpha I)^{-1}y \tag{6.8}$$

The direct and indirect forms for solving the inverse problem also arise in blood oxygenation imaging using multispectral diffuse optical tomography (DOT) after linearization of the DOT forward problem. The multispectral DOT problem is mathematically identical to MSFT as de-

tailed above, except that the oxy and deoxy-hemoglobin absorption spectra replace the fluorescence emission spectra in (6.1)-(6.6). Although a systematic comparison of indirect and direct approaches for MSFT has not been conducted previously, properties of both approaches have been analyzed in multispectral DOT. These results will be summarized in the next section.

For ease of comparison of TD and spectral methods for multiplexing, we restate the direct TD (DTD) and asymptotic TD (ATD) forward and inverse operators below:

$$\begin{array}{ll} DTD: y = W\eta & \qquad \eta^{DTD} = W^T(WW^T + \lambda I)^{-1} y \\ ATD: y = A\overline{W}\eta & \Longleftrightarrow \qquad \eta^{ATD} = \overline{W}^T(\overline{W}\overline{W}^T + \lambda I)^{-1} A^{\dagger} y \end{array} \tag{6.9}$$

It can be seen that the DTD approach is analogous to the direct approach in the spectral domain while the ATD approach is analogous to the indirect approach. However, it is clear from a comparison of (6.5) and (6.9) that there exists a fundamental difference between the spectral and TD forward problems. Although both the full TD and spectral weight matrices can be written as a product of a block diagonal matrix containing CW weight matrices and basis function matrix containing either the excitation/emission spectra (for the spectral case) or the exponential decay functions (for the TD case), the order of these two matrices in the product is different for the multispectral and lifetime problems. In the case of spectral tomography, the mixing of the unknown fluorophore concentrations occurs at the level of the individual voxels. The mixed concentrations are then propagated through the medium by wavelength dependent CW weight matrices. In the TD case, the separated unknown concentrations are propagated through the medium by reduced absorption CW matrices first and then mixing occurs at the level of the measurement. This difference in ordering results in a fundamental difference between the capability of each method to separate fluorophores.

To evaluate the separation capability of spectral compared to TD, we measure the amount of

cross-talk that results for both methods. In chapter 4, we showed that cross-talk in tomographic reconstructions of fluorophores can result in error in quantifying relative amounts of multiple fluorophores and inaccurate localization for closely spaced fluorescent objects. We also showed that a general way to evaluate cross-talk throughout an imaging medium for a linear inverse operator is to consider the values of the off-diagonal blocks of the resolution matrix. For the ATD method, we showed the resolution matrix reduces to a block diagonal matrix due to cancellation of the basis matrix, A with its pseudoinverse, $A^{\dagger}$:

$$\begin{aligned} R_{ATD} &= W_{ATD}^{-1} W \\ &= \overline{W}^T (\overline{W}\overline{W}^T + \alpha I)^{-1} A^{\dagger} (A\overline{W}) \end{aligned} \tag{6.10}$$

However, if we consider the resolution matrix for the two spectral methods, we see that the resolution matrix is not generally block diagonal for either method. In particular, for the indirect method we see that due to reversal in ordering of the factorization in (6.5), the basis matrix $A^{\dagger}$ is no longer canceled by its pseudoinverse as occurs for ATD:

$$\begin{aligned} R_{Direct} &= W_{Direct}^{-1} W \\ &= W^T (WW^T + \alpha I)^{-1} W \end{aligned} \tag{6.11}$$

$$\begin{aligned} R_{Indirect} &= W_{Indirect}^{-1} W \\ &= A^{\dagger} \overline{W}^T (\overline{W}\overline{W}^T + \alpha I)^{-1} (\overline{W} A) \end{aligned} \tag{6.12}$$

From (6.11), it can be seen that $R_{Indirect}$ is only block diagonal when either or both of the following conditions are true:

- The spectra for the different flouorophores are orthogonal to each other. This is equivalent to completely non-overlapping excitation and emission spectra.

- The individual CW weight matrices along the block diagonal in $\overline{W}$ are identical (i.e. $\overline{W_i} = \overline{W_j} = \overline{W_0}$ in (6.6), so that $\overline{W} = \overline{W_0} * I$). This is equivalent to the condition that optical properties of the medium do not change with wavelength.

Both the above conditions are hard to satisfy for practical fluorophores and tissue optical properties. It is interesting to compare the resolution matrix for ATD inversion (6.9) which is block diagonal under the well satisfied condition that $A^{\dagger}A = 1$. Thus, lifetime multiplexing using the ATD approach is fundamentally better posed for un-mixing multiple fluorophores as compared to spectral un-mixing. In what follows, we will demonstrate the difference between TD and spectral methods using simulations with multiple infrared fluorescent proteins (iRFP) in diffuse media.

6.2 Multispectral Diffuse Optical Tomography

Before simulation results are presented for MSFT, we give a brief overview of reconstruction of oxy and deoxy-hemoglobin using multispectral DOT where both direct and indirect forms have been used and comparisons of the two methods have been performed. We know from chapter 2 that in the DOT forward problem, the measurements are a nonlinear function of μ_a and μ_s. For imaging of dynamic changes in blood oxygenation, it is generally assumed changes in scattering is negligible and any change in measurements results from a change in absorption. One approach to solve the nonlinear inverse problem is to linearize about a particular set of μ_a values. Using this approach, the Jacobian, J_a relating a change in measurement Δy to a change in absorption, $\Delta\mu_a$ can be calculated using the Born approximation. The change in absorption can then be spectrally divided into ΔHbO and ΔHbR using a Jacobian matrix J_s which contains the extinction coefficent spectra for both chromophores. The Jacobian that directly relates Δy to ΔHbO and ΔHbR,

J_d is the product of J_a and J_s. As in the case for MSFT, the inversion can either be performed with the full weight matrix (direct) or in two steps (indirect). Both forms are shown below:

$$\begin{bmatrix} \Delta HbO \\ \Delta HbR \end{bmatrix} = J_d^T (J_d J_d^T + \alpha I)^{-1} \Delta y \tag{6.13}$$

$$\begin{bmatrix} \Delta HbO \\ \Delta HbR \end{bmatrix} = J_s^{\dagger} J_a^T (J_d J_a^T + \alpha I)^{-1} \Delta y \tag{6.14}$$

A key difference between multispectral DOT and MSFT is that typically measurements are only collected at two wavelengths for multispectral DOT while multiple wavelengths are collected for MSFT. When comparing cross-talk performance of both methods, (Li et al., 2004) showed that cross-talk is higher for the direct approach than the indirect approach for two sets of wavelengths. On the other hand, (Zhan et al., 2012) showed with simulations that the indirect approach had lower cross-talk than the direct approach. The discrepancy may lie in how regularization was chosen for each study. (Li et al., 2004) regularized both methods such that the full width at half maximum (FWHM) of the reconstructions were the same. This may result in under regularization of the indirect approach which could lead to increased cross-talk. We have found that the regularization parameters which produce the smallest reconstruction error results in significantly different FVHM's for direct and indirect methods. In our simulations, cross-talk of indirect and direct approaches for MSFT will be compared using the off diagonal elements of their resolution matrices. These results should also be generalizable to multispectral DOT.

6.3 Infrared Fluorescent Proteins

Whole body imaging of fluorescent proteins in mice would allow for non-invasive monitoring of biological processes such as gene expression without the need for external contrast agents. Most current studies use fluorescent proteins in the visible to far red regions of the spectrum where due to absorption by hemoglobin and melanin, the penetration depth is low and the background signal from tissue autofluorescence is high. Measurements are generally collected in the reflection geometry and reconstruction is limited to regions close to the mouse surface. A new set of near infra-red fluorescent proteins (iRFP) with excitation and emission in the near infrared have the potential to allow imaging throughout the mouse body. Their unique spectra and lifetimes should allow for multiplexing of multiple processes in vivo. We use these iRFP's as a model to compare the fluorophore separation capability of TD and spectral methods.

The lifetimes of three iRFP's (iRFP 670, 702 and 720) from (Shcherbakova & Verkhusha, 2013) have been characterized in bacteria. Reflectance measurements were performed with excitation at 700 nm and with emission collected with a 750 nm long pass filter. Fig. 6.1 (b) shows the measured decay functions for the three bacteria streaks. Histograms for the three iRFP's showed narrow lifetime distributions with lifetimes centered at $\tau = 0.68$, 0.78 and 0.93 ns. Fig. 6.1 (a) shows plots of the excitation and emission spectra taken from (Shcherbakova & Verkhusha, 2013).

6.4 Simulation Parameters

All forward simulations were performed with the photon transport software package, tMCimg. Spectral simulations used eight wavelengths (650 nm to 790 nm in steps of 20 nm) from each iRFP's emission spectrum. Time domain simulations used 30 times gates in the asymptotic regime.

Two imaging mediums were considered in this study: a rectangular slab and a digital mouse atlas.

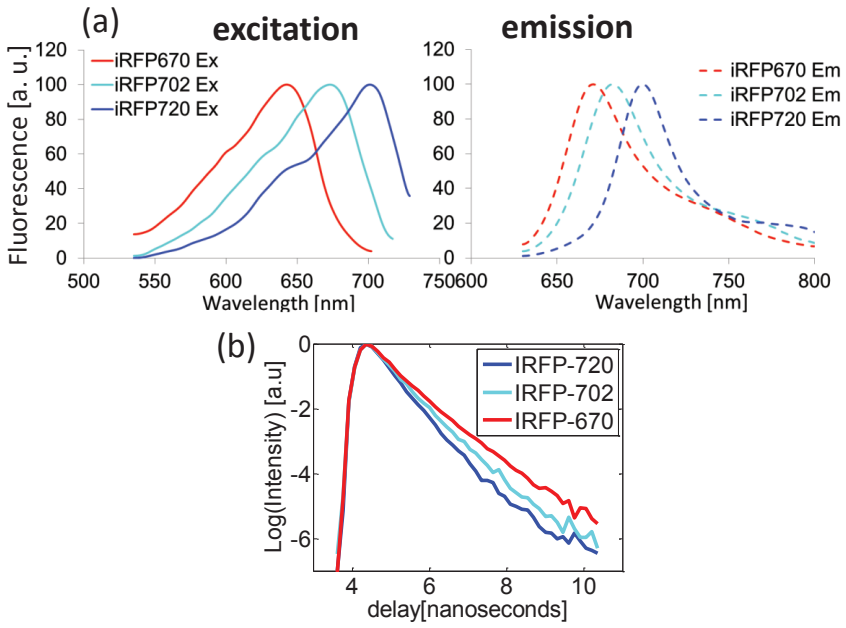

Figure 6.1: (a) Excitation (left) and emission (right) spectra for the three iRFP's (iRFP 670, 702 and 720) used in this study (taken from (Shcherbakova & Verkhusha, 2013)). (b) Measured fluorescence decay functions for streaks of bacteria expressing the three iRFP's.

The rectangular slab was used for simulation results presented in sections 6.5.1, 6.5.2 and 6.5.3. The slab size was 2 cm $\times$ 2 cm $\times$ 2 cm with 1 mm^3 voxel size. 49 sources and 49 detectors were located at the z = 0 and z = 2 cm planes respectively. The background absorption coefficient of the medium was varied linearly as a function of wavelength while the background reduced scattering coefficient was held constant at $\mu_s' = 10\,\mathrm{cm}^{-1}$.

The separation of three iRFP's was investigated in a realistic mouse model (Dogdas et al., 2007) (results in section 6.5.4). The mouse atlas was first binned down from 0.1 mm voxel length to 1 mm voxel length. 72 sources were placed under a region of the torso covering parts of the skeleton, lung, heart, liver and kidneys. An equal number of detectors were located directly above each source at the top surface of the mouse model. Fig. 6.2 shows the coverage of the source/detectors and the included organs highlighted with different colors. Background absorption values were assumed to depend only on the absorption coefficients of oxy and deoxy-hemoglobin and water (μ_{aHbO}, μ_{aHbR} and μ_{aW}). The total absorption in an organ can be modeled as a weighted sum of each chromophore using:

$$\mu_a(\lambda) = S_B(x\mu_{aHbR}(\lambda) + (1-x)\mu_{aHbO}(\lambda)) + S_W\mu_{aW}(\lambda) \tag{6.15}$$

where S_B and S_W are scaling factors for blood and water components and x is the oxygen saturation, SO_2.

The scattering spectra for each organ was assumed to follow an inverse power law:

$$\mu_s'(\lambda) = a \times \lambda^{-b}, \;\; \lambda \text{ in nm} \tag{6.16}$$

where a is the scattering amplitude and b is the scattering power.

In (Alexandrakis et al., 2005), the parameters S_B, S_W, x, a and b in (6.15) and (6.16) were em-

Tissue type	S_B	x	S_W	$a(mm^{-1})$	b
Bone	0.049	0.8	0.15	35600	1.47
Lung	0.15	0.85	0.85	68.4	0.53
Liver	0.30	0.75	0.7	629	1.05
Heart	0.30	0.75	0.7	629	1.05
Kidneys	0.056	0.75	0.8	41700	1.51
Muscle	0.07	0.8	0.5	4e7	2.82

Table 6.1: Optical parameters S_B, S_W, x, a and b used in (6.15) and (6.16) to determine $\mu_a(\lambda)$ and $\mu_s(\lambda)$ for each segmented organ. Tables values are taken from (Alexandrakis et al., 2005).

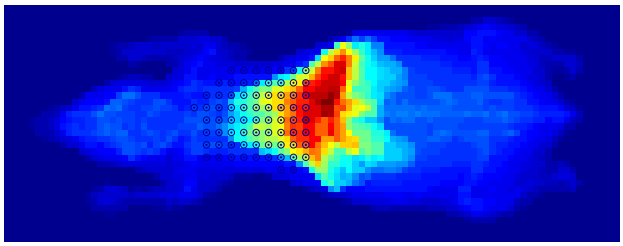

Figure 6.2: Horizontal view of Digimouse model (Dogdas et al., 2007) with source/detector positions labeled as black circles.

pirically fit using data from literature for various organs. Table 6.1 summarizes the parameter values for the organs used in this study. The absorption spectra of individual chromophores (μ_{aHbR}, μ_{aHbO}, μ_{aW}) were obtained from (Prahl, 2001). Since parameters were not available for the heart, the parameters for the liver was used in its place as both are highly absorbing organs. Voxels in the mouse atlas not corresponding to any of the five organs was modeled using muscle optical properties.

6.5 Results

6.5.1 Resolution Matrix Comparison

The resolution matrices for ATD, multispectral indirect and direct approaches, R_{ATD}, $R_{Indirect}$ and R_{Direct} were computed using (6.10)-(6.12). For spectral methods, the absorption coefficient was assumed to increase linearly by 100% over the wavelength range. For ATD, the absorption coefficient was taken to be absorption at the first spectral wavelength. The emission spectra and lifetimes of iRFP702 and iRFP720 were used for multiplexing. The cross-talk is represented by any terms in the off diagonal blocks of the resolution matrix. For aid in visualization, the rows and columns of resolution matrices were both binned by a factor of 40. Fig. 6.3 shows the resolution matrices for all three methods. It can be seen that the crosstalk terms for R_{Direct} are positive and generally symmetric between the two iRFP's. $R_{Indirect}$ shows more asymmetry between the two iRFP's with a major portion of the cross-talk being negative. As expected, all cross-talk terms for R_{ATD} are zero since the ATD resolution matrix is block diagonal. To quantify cross-talk between direct and indirect spectral methods, cross-talk was evaluated at each voxel by taking the absolute value of the ratio between the on-diagonal and off-diagonal terms of the resolution matrix corresponding to that voxel. Line plots showing cross-talk for both iRFP's are shown in Fig. 6.3. It can be seen that cross-talk in the indirect approach is generally lower than that for the direct approach.

6.5.2 Cross-talk vs. Optical Property Change

It is known from (6.12) that for the indirect method, a change in optical properties at different wavelengths results in cross-talk and a flat optical property spectrum would result in zero cross-talk. We look to explore the relationship between cross-talk and the variation of the optical property spectrum. Different optical property spectra are created by taking a particular starting bulk

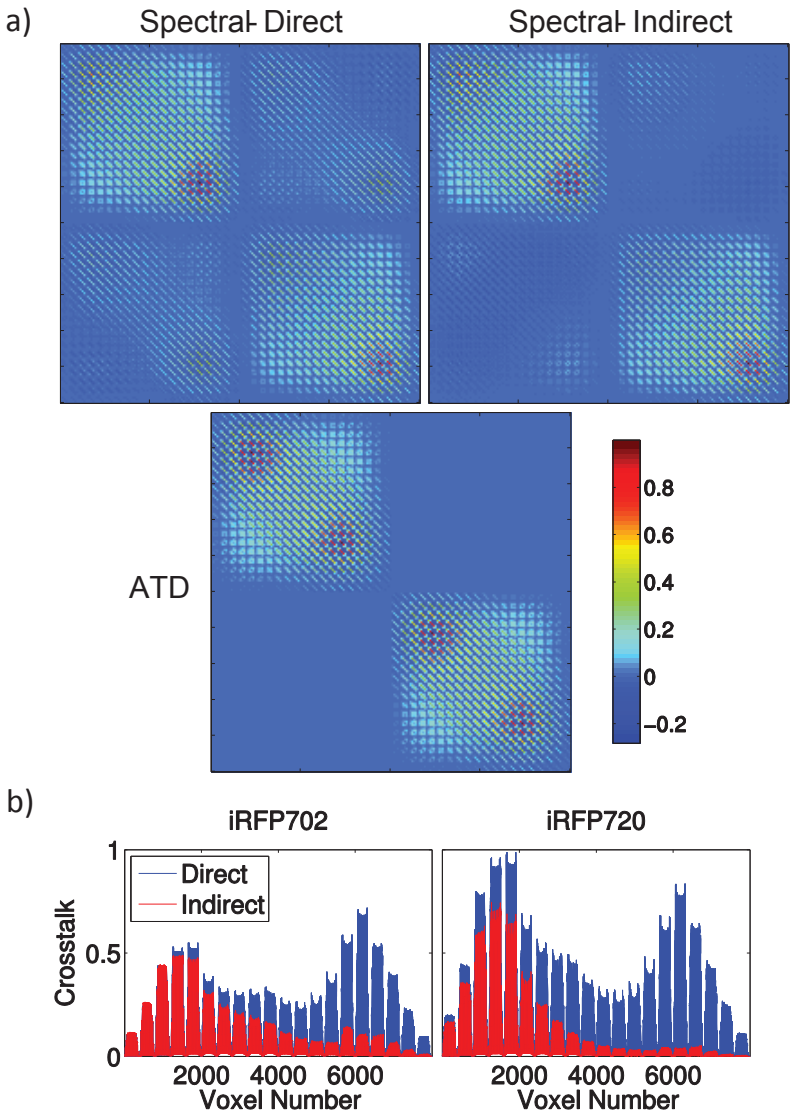

Figure 6.3: (a) Resolution matrices for spectral methods (indirect and direct) and ATD. (b) The cross-talk for overlapping fluorophores is plotted at each voxel of the medium for direct (blue) and indirect (red) methods.

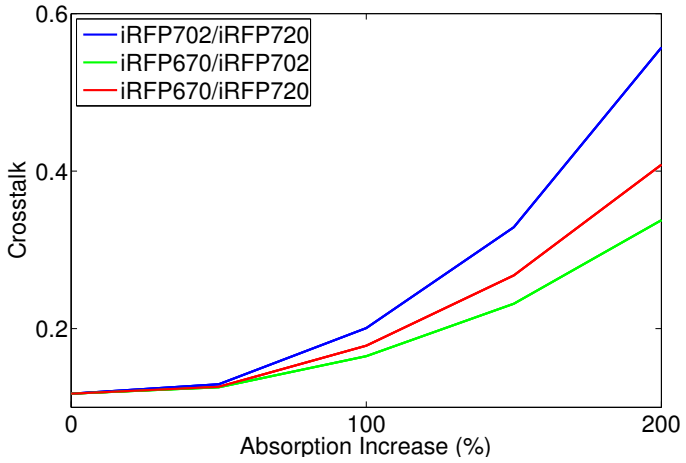

Figure 6.4: Total cross-talk for different amounts of variation of the optical property spectra. Optical property spectra are increased linearly from a starting absorption of $\mu_a = 0.6\,\mathrm{cm}^{-1}$.

absorption ($\mu_a = 0.6\,\mathrm{cm}^{-1}$) and linearly increasing it by different amounts over the range of wavelengths. The total cross-talk is evaluated for each of these optical property spectra. From Fig. 6.4, it can be observed that cross-talk increases with larger variations of the absorption spectra for the three pairs of iRFP's.

6.5.3 Cross-talk vs. Error in Optical Properties

In the previous section, the cross-talk was compared for various changes in optical properties when they were exactly known at each wavelength. However, the optical property distribution in a living mouse is widely varying and with current techniques for in vivo optical property reconstruction, only low resolution reconstructions with large errors can be achieved. We investigated the effect of error in the model optical properties on cross-talk performance. Spectral data was generated assuming an absorption spectra that increased by 100% over the iRFP emission wavelengths. Error optical property spectra were generated using the true spectra, ϵ_{true} and a flat

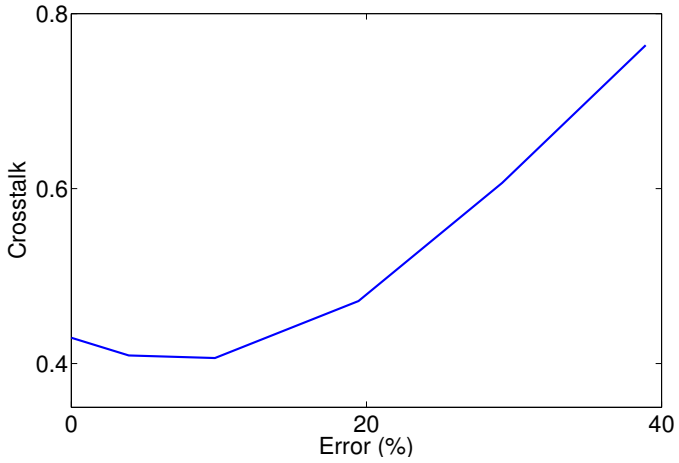

Figure 6.5: Total cross-talk for different amounts of error in the model optical properties of the medium. Optical property spectra was increased by 100% over the wavelength range of the spectral measurements.

spectra, ϵ_{flat}:

$$\epsilon_{error} = \epsilon_{true} + f(\epsilon_{flat} - \epsilon_{true}), 0 \leq f \leq 1 \tag{6.17}$$

For the indirect approach, cross-talk was found to increase significantly as the error in the model optical properties was increased (Fig. 6.5).

6.5.4 Multiplexing of Three iRFP's in Mouse Atlas

To demonstrate the difference in separation capability for TD and spectral methods, we consider the simple problem of determining the relative amounts of three iRFP's embedded within a mouse. The imaging medium was a realistic mouse model with optical property spectra derived from literature. A single voxel sized fluorescent inclusion was placed at the center of the heart containing equal amounts of all three iRFP's. Both ATD and spectral methods were used to reconstruct for the inclusion. Fig. 6.6 shows the recovered concentration for each iRFP normalized to the concentration of iRFP720. It is clear that the ATD method could accurately recover the relative amounts of each iRFP as recovered concentrations did not differ by more than 3.7%

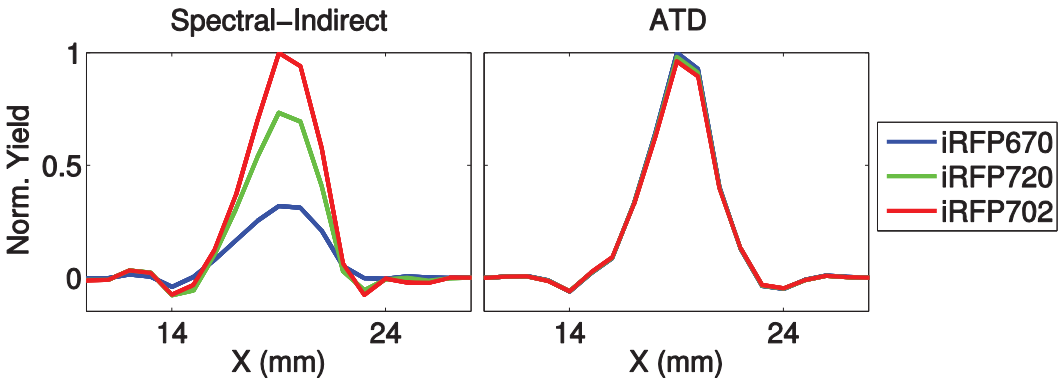

Figure 6.6: Reconstruction of three colocalized iRFP's using indirect-spectral (left) and ATD (right). The normalized yields for iRFP702, iRFP720 and iRFP670 are shown in red, green and blue respectively.

from each other. However, due to cross-talk the indirect spectral method could not accurately recover the correct relative amounts. The yield for iRFP702 differed from iRFP720 by 27.7% and iRFP670 differed from iRFP720 by 68.1%.

To verify that cross-talk is the source of this error in quantitation, we performed spectral reconstruction in the two cases when cross-talk could be eliminated for spectral methods. In the first case, the iRFP basis functions were made orthogonal by setting two of the three emission spectra to zero at each wavelength. In the second case, the optical property spectrum was kept constant for all wavelengths. Reconstructions in both cases showed accurate recovery of the relative concentrations for each iRFP.

6.6 Discussion

We have used recently developed iRFP's to compare the performance of separating multiple fluorophores with TD and spectral methods. In a mouse atlas model, we have shown the ability of ATD to accurately recover the relative amounts of three iRFP's while cross-talk prevented spectral methods from obtaining an accurate estimate.

For a general problem of separating multiple fluorophores, the amount of cross-talk and degree of quantitation error for spectral methods will depend on many factors. In this study, we have shown the dependence of cross-talk on the fluorescence spectra, the variation in optical properties over the wavelength range of interest and the accuracy of the optical properties used to model the medium. Other factors that may also influence separability include the number of wavelengths used and noise level of the system. However, the central result of this study remains true. TD methods provide an inherent advantage for separating fluorophores in deep tissue because it allows un-mixing of the fluorophores at the measurement level while in spectral methods measurements are distorted by propagation through a diffuse medium and un-mixing of the fluorophores occurs at the voxel level after attempting to correct for this distortion.

The effect of optical property spectra on multispectral reconstructions has also been investigated for single fluorophore reconstructions (Swartling et al., 2005; Chaudhari et al., 2009; Leblond et al., 2011). In these studies, multispectral data has been used to determine the depth of a fluorescent inclusion and it has been found that large changes in the optical properties at different wavelengths improves the localization due to reduced redundancy in the data. Since the spectrum is more slowly varying for longer near-infrared wavelengths, excitation-resolved tomography was shown to perform better than emission-resolved tomography. Our results show the opposite is true when the goal is to separate multiple fluorophores. A flatter spectrum for optical properties results in less distortion of the fluorophore spectra by the medium. This ultimately leads to less cross-talk and better quantitation of relative amounts of each fluorophore.

7

Constrained Optimization in Optical Tomography

7.1 Introduction

In chapter 4, we have shown that in fluorescence lifetime tomography, the TD measurements can be inverted in two ways. From the perspective of a Bayesian framework and assuming white

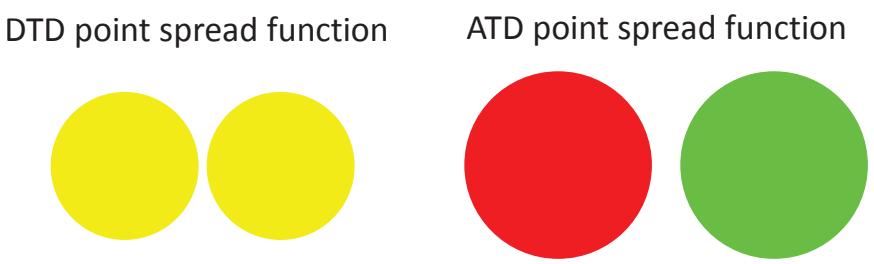

Figure 7.1: Representative responses to two point inclusions with different lifetimes for DTD (left) and ATD (right). Red and green represent the individual yield components of each fluorophore while yellow represents cross-talk.

Gaussian processes for the measurement and fluorescence yields, the Direct TD (DTD) solution is equivalent to the minimum mean square error (MMSE) solution for appropriate choice of the regularization parameter. We saw that by replacing the actual measurement covariance matrix with a matrix formed from exponential basis functions, we could obtain another estimator (ATD approach) which results in better separation of the fluorophores at the cost of mean square error. The cartoon in Fig. 7.1 illustrates the features of both estimators. Red and green represent the yields components of individual fluorophores and the size of the circle represents the point spread function or resolution of the method. ATD has larger circles but with complete separation of the fluorophore while DTD has smaller circles with the yellow in each circle representing a mixture of the two fluorophores due to cross-talk.

In this chapter, we seek to obtain new estimators which provide better separation than standard estimators in more general multiple parameters separation problems. We know that the ATD method depends on the factorization of the full TD weight matrix into a well conditioned temporal basis matrix and a spatially dependent matrix containing CW weight matrices. For many problems in TD optical tomography, this factorization is not possible and a new approach is needed. Two problems in optical tomography where a general parameter separation algorithm can be applied are absorption and scattering separation in diffuse optical tomography and the separation of fluorophores using TD data not within the asymptotic regime where an ATD reconstruction

method cannot be applied.

7.2 Absorption and Scattering Separation

We recall from chapter 2 that a nonlinear function relates the experimental measurements at the surface of a diffuse medium, y with the absorption coefficient, $\mu_a(r)$ and scattering coefficient, $\mu_s(r)$ throughout the medium:

$$y = \Phi(\mu_a(r), \mu_s(r)) \tag{7.1}$$

Solving the inverse problems requires linearization of Φ to find the absorption and scattering Jacobians (J_{μ_a} and J_{μ_s}) about a set of background optical properties, $\mu_{ao}(r)$ and $\mu_{so}(r)$. The Jacobians are used to relate the difference between the unknown and background optical properties ($\Delta\mu_a = \mu_a - \mu_{ao}$, $\Delta\mu_s = \mu_s - \mu_{so}$) and the difference between the experimental measurement and the simulated measurement through the background medium ($\Delta y = y - \Phi(\mu_{ao}, \mu_{so})$):

$$\Delta y = \begin{bmatrix} J_{\mu_a} & J_{\mu_s} \end{bmatrix} \begin{bmatrix} \Delta\mu_a \\ \Delta\mu_s \end{bmatrix} \tag{7.2}$$

For simultaneous imaging of the absorption and scattering coefficient, (Arridge & Lionheart, 1998) has shown that in a highly scattering medium where the diffusion approximation holds, it is generally not possible to uniquely determine both using CW data. Uniqueness of absorption and scattering requires time-resolved data types such as frequency domain (FD) or time-domain (TD). To illustrate this point, Fig. 7.2 shows the simulated response of a pure absorption and pure scattering inclusion for the two measurement types. For the spatial response of CW in Fig. 7.2(b, d) it can be seen that the normalized δy for absorption and scattering are nearly indistinguishable. On the other hand, the normalized δy for the TD response in Fig. 7.2(c) shows distinct behavior between the absorption and scattering responses with the scattering temporal profile preceding the

absorption temporal profile.

Despite the ability of TD methods to better distinguish μ_a and μ_s' than CW, there still remains a large cross-talk component when simultaneously reconstructing for both parameters (Gao et al., 2002). To bypass this problem many studies have assumed pure absorption or scattering perturbations. There have also been attempts to use differences between the temporal profile of the absorption and scattering response to provide better separation. In (Nouizi et al., 2011), a reconstruction algorithm was devised using points near the tail of the TPSF to reconstruct solely for absorption while points on the rising portion were used to reconstruct solely for scattering. It was found that cross-talk between absorption and scattering was reduced compared to reconstructions using the full TD data. An alternate approach was introduced in (Painchaud et al., 1999) which utilized temporal responses to single absorption and scattering inclusions at the center of a scattering slab as basis functions. A linear fit of the TD measurement data to the basis functions was able to discriminate absorption and scattering inclusions in the center plane while performance was poor for other regions in the medium. In this chapter, we seek a general method which can be applied for inclusions throughout the medium and which optimally uses all time points on the TPSF to separate both absorption and scattering simultaneously. To achieve this aim, we investigate using a constrained optimization approach for parameter separation in the next section.

7.3 Constrained Mean Square Error Estimator

In this section, we introduce a new reconstruction method based on the Bayesian approach for solving linear inverse problems. A general linear inverse problem with a linear estimation matrix,

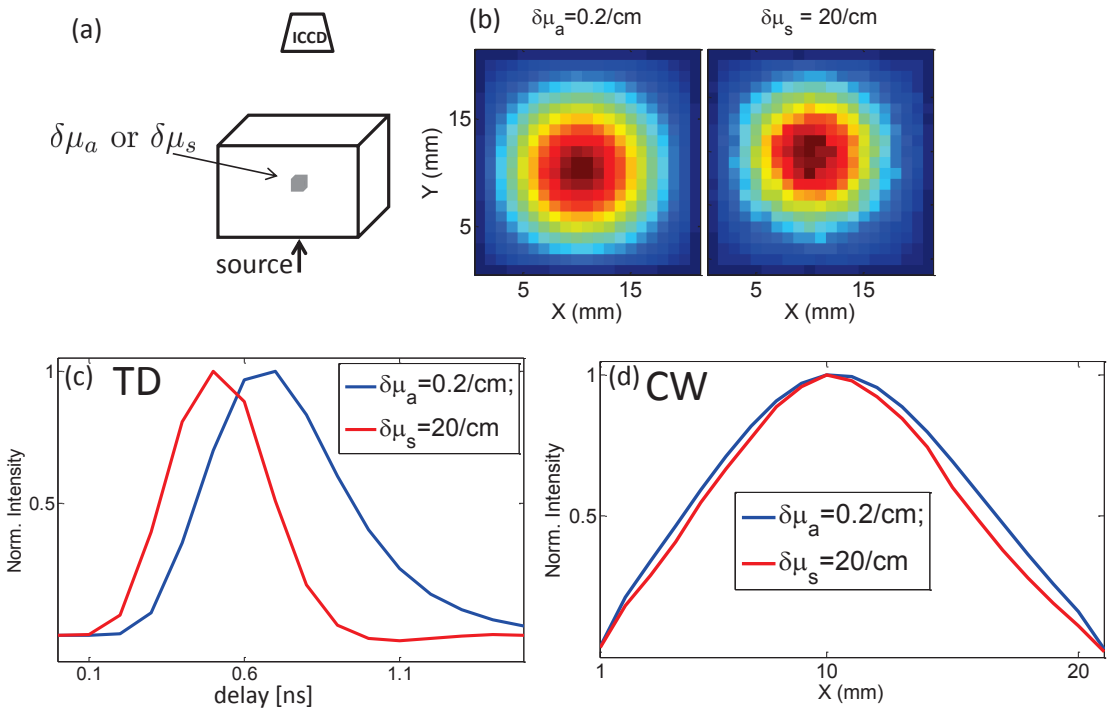

Figure 7.2: A comparison of CW and TD surface measurements for absorption and scattering inclusions. (a) The perturbed measurement due to $\delta\mu_a = 0.2\,\text{cm}^{-1}$ and $\delta\mu_s' = 20\,\text{cm}^{-1}$ inclusions are convolved with the instrumental response function of our TD system. (b) Spatial response due to absorption and scattering inclusions. (c) Normalized temporal response of absorption (blue) and scattering (red) inclusions. (d) Line plot of the spatial response due to absorption and scattering inclusions.

W can be described as follows:

$$y = Ax_{true} + n \tag{7.3}$$

$$x_{recon} = Wy \tag{7.4}$$

It is known that the linear minimum mean square estimator (LMMSE) can be derived by minimization of the following functional:

$$W_{MMSE} = \arg\min_{W} E\left[\|x_{recon} - x_{true}\|^2\right] \tag{7.5}$$

The solution of (7.5) can be obtained by differentiation with respect to W. Under the assumption that the noise n is zero mean with covariance matrix, C_n, the object x_{true} is zero mean with covariance matrix, C_x and n and x_{true} are independent random vectors, this minimization leads to the LMMSE estimator:

$$W_{MMSE} = C_x A^T (AC_x A^T + C_n)^{-1} \tag{7.6}$$

An alternate form which is mathematically equivalent to (7.6) but is more computationally efficient in the case when number of unknowns is greater than number of measurements is given as:

$$W_{MMSE} = (A^T C_n^{-1} A + C_x^{-1})^{-1} A^T C_n^{-1} \tag{7.7}$$

It should be noted that the LMMSE only depends on the second order statistics of n and x_{true} and does not assume them to be Gaussian. However, the minimum mean square error (MMSE) estimator is the same as the LMMSE when n and x_{true} are jointly Gaussian. To see the connection between the Bayesian approach and Tikhonov regularization, we consider the form of (7.6) and (7.7) when n and x_{true} are white processes with $C_n = \sigma_n^2 I$ and $C_x = \sigma_x^2 I$. In this case, the LMMSE

estimator is the same as the Tikhonov regularization inverse operator with the choice of regularization parameter, $\lambda = \sigma_n^2/\sigma_x^2$.

We have seen for the tomographic lifetime multiplexing problem, the linear MMSE estimator does not necessarily correspond to the linear estimator which optimizes for other performance measures such as inter-parameter cross-talk. We look to derive a new estimator for a multiple parameter inverse problems which minimizes for error while placing a restriction on cross-talk. As a first step, we define cross-talk in a statistical manner.

We know cross-talk arises during reconstruction in a multiple parameter problem when the presence of one parameter type in a spatial location results in reconstruction of other parameter types in the same location. This leads to a natural definition for cross-talk as the correlation between the true object of a particular parameter type with the reconstructed objects of other parameter types. We consider the case of two parameters where $x_{true} = [(x_{1\_true})^T (x_{2\_true})^T]^T$ and $x_{recon} = [(x_{1\_recon})^T (x_{2\_recon})^T]^T$. The cross-talk for x_1 and x_2 can be expressed in terms of their cross-talk matrices, $Crosstalk_1$ and $Crosstalk_2$:

$$Crosstalk_1 = E\left[(x_{1\_true})\,(x_{2\_recon})^T\right] \tag{7.8}$$

$$Crosstalk_2 = E\left[(x_{2\_true})\,(x_{1\_recon})^T\right] \tag{7.9}$$

Element ij of the cross-talk matrices represents the response at voxel j of parameter 1 (2) due to a true object at voxel i of parameter 2 (1). We see that $E[(x_{true})\,(x_{recon})^T]$ can also be expressed in terms of the resolution matrix $R = WA$:

$$E\left[(x_{true})\,(x_{recon})^T\right] = E[x_{true}(WAx_{true} + Wn)^T = C_x(WA)^T \tag{7.10}$$

Therefore, from (7.10), our statistical definition of cross-talk also corresponds to the off-diagonal

blocks of the resolution matrix when C_x is the identity matrix.

We now define a new estimator based on minimizing the mean squared error while imposing linear constraints that all terms of the cross-talk matrices must be zero using the method of Lagrange multipliers. We combine the MMSE functional (7.5) with linear equality constraints on (7.8) and (7.9) to form the Lagrangian below:

$$L(W, \lambda 1_{ij}, \lambda 2_{ij}) = E\left[\|x_{recon} - x_{true}\|^2\right] + \sum_{i=1}^{V}\sum_{j=1}^{V} \lambda 1_{ij} E\left[\left(x_{1\_true}\right)\left(x_{2\_recon}\right)^T\right]_{ij} + \sum_{i=1}^{V}\sum_{j=1}^{V} \lambda 2_{ij} E\left[\left(x_{2\_true}\right)\left(x_{1\_recon}\right)^T\right]_{ij} \quad (7.11)$$

where $\lambda 1_{ij}$ and $\lambda 2_{ij}$ are Lagrange multipliers and V is the number of voxels in the medium. The W which minimizes (7.11) while satisfying the linear constraints can be found by solving:

$$\nabla L = 0 \quad (7.12)$$

Although we have shown a general method for minimizing cross-talk using a linear estimator, an important issue to consider is the increase in error due to application of constraints. The constraints that are being imposed in (7.11) are effectively shaping the resolution matrix to be closer to a diagonal matrix. We know from chapter 2 that as the resolution matrix approaches the identity matrix, the approximation error decreases while the noise amplification error increases. Since any constraint on the LMMSE functional will result in sub-optimal solutions, the total MSE will also increase relative to the LMMSE solution. When too many constraints are imposed or the constraints are too restrictive, the gain in cross-talk performance may not be worth the loss in error performance. Preliminary simulations show that the noise amplification error is much above the reconstructed objects when a constraint is placed on every term of the cross-talk matrix. How-

ever, a compromise between cross-talk performance and keeping a low noise amplification error can be achieved using fewer constraints. Since the highest cross-talk term for parameter n at voxel i is frequently also at voxel i for other parameters, we have found that estimators with only constraints on these voxels can perform well in reducing cross-talk without an increase in noise level that will overwhelm the solution. For the two parameter case, this corresponds to constraints on the diagonal elements of the cross-talk matrices:

$$E\left[\left(x_{1\_true}\right)\left(x_{2\_recon}\right)^T\right]_{ii} = 0, \text{ for } 1 \leq i \leq V \tag{7.13}$$

$$E\left[\left(x_{2\_true}\right)\left(x_{1\_recon}\right)^T\right]_{ii} = 0, \text{ for } 1 \leq i \leq V \tag{7.14}$$

Also, despite only placing a single constraint for each voxel we have found that after solving for the optimal constrained W, the resulting cross-talk matrix has many of its terms reduced. The estimator using the constraints in (7.13) and (7.14) are used for all simulations and experimental results presented in this chapter. In later sections, we investigate this estimator for two types of multiple parameter problems in optical tomography: separation of absorption and scattering in DOT and separation of short lifetime components which do not satisfy the asymptotic condition in fluorescence tomography.

7.4 Diffuse Optical Tomography Jacobian Matrices

Before optical property reconstructions can be performed, a method to generate time dependent Jacobian matrices for absorption and scattering have to be implemented. Previous optical property work in our lab have used frequency domain reconstructions of Fourier transformed time-domain data where the Jacobian matrices were constructed using adjoint solutions to the diffusion equation. To ensure accuracy of our optical reconstructions even for conditions when the diffusion

equation may not hold such as early time gates, we look to implement time domain Jacobian matrices using the radiative transport equation (RTE). Two methods were implemented in this study: an adjoint approach based on the RTE and a perturbation Monte Carlo (pMC) approach. In each approach, an open source photon propagation Monte Carlo software, tMCimg (Boas et al., 2002) was modified to allow for Jacobian calculation.

7.4.1 Transport based adjoint method

The adjoint form for the RTE can be derived in the following way. First, the parameters and fluence are written as a background component and a differential component:

$$\mu_a = \mu_{a0} + \delta\mu_a \tag{7.15}$$

$$\mu_s = \mu_{s0} + \delta\mu_s \tag{7.16}$$

$$G = G + \delta G^a + \delta G^s \tag{7.17}$$

Substituting of (7.15)-(7.17) into the RTE and neglecting first order terms results in:

$$\delta G^a(r_s, r_d, \hat{s}, t) = -\int d^3 G(r_d, r', t, \hat{s}) \bigotimes G(r', r_s, t, \hat{s})\delta\mu_a(r) \tag{7.18}$$

$$\delta G^s(r_s, r_d, \hat{s}, t) = -\int d^3 rG(r_d, r', \hat{s}, t) \bigotimes Q(r', r_s, \hat{s}, t)\delta\mu_a(r) \tag{7.19}$$

$$Q(r, r_s, \hat{s}, t) = G(r, r_s, \hat{s}, t) - \int d\hat{s}' P(\hat{s} \cdot \hat{s}')G(r, r_s, \hat{s}', t) \tag{7.20}$$

When compared to the Jacobian matrices derived for the diffusion equation (chapter 2), it can be seen that the form for the absorption Jacobian for both cases is very similar. The main difference is that in the transport based formulation, there is an additional dependence on $\hat{s}$. On the other hand, the scattering Jacobians for transport and diffusion equations have quite different forms.

While the scattering Jacobian for the diffusion equation requires the computation of spatial gradients and a temporal convolution, the transport based method is calculated from the difference between the source Green's function and the source Green's function weighted by the scattering phase function. It was observed that the spatial gradient term of the diffusion equation based scattering Jacobian produced artifacts near boundaries and other sharp transitions which were not present in the transport based scattering Jacobian.

In our current Monte Carlo implementation, the fluence due to a point source is stored at each spatial location and for each time gate. This quantity is equivalent to the volume Green's function. From (7.20), it can be seen that Jacobian matrix calculation using the transport based adjoint methods requires additional storage of the directions of every simulated photon. tMCimg was modified by creation of azimuthal and deflection angle variables and assigning the photon's direction to angle bins after every scattering event. Although the forms for the Jacobian in (7.18) and (7.19) were reported in (Arridge, 1999) and a CW implementation was used in (Kim et al., 2006), to our knowledge, time domain version has not previously been implemented.

7.4.2 Perturbation Monte Carlo Method

An alternate method for calculation of absorption and scattering Jacobians is based on the perturbation Monte Carlo (pMC) approach (Hayakawa et al., 2001; Sassaroli et al., 1998; Chen & Intes, 2009). Using the photon path histories for an unperturbed medium, the pMC approach can predict the effect on the measurement from a perturbing region. The strength of the pMC approach is its the ability to calculate this gradient information from a single Monte Carlo simulation compared to standard methods which would require multiple Monte Carlo simulations. (Hayakawa et al., 2001) showed that the measurement w due to a single photon in a medium with

a perturbation can be expressed as:

$$w = w_0 \left(\frac{\hat{\mu}_s / \hat{\mu}_t}{\mu_s / \mu_t} \right)^k \left(\frac{\hat{\mu}_t}{\mu_t} \right)^j exp\left[-(\mu_t^* - \mu_t) l \right] \tag{7.21}$$

where w_0 is the measurement in a medium with no perturbations, k is the number scattering events in the perturbed region, l is the total pathlength inside the perturbed region and the optical properties in the perturbed region are given by: $\hat{\mu_s} = \mu_s + \Delta\mu_s$ and $\hat{\mu_a} = \mu_a + \Delta\mu_a$.

When (7.21) is extended to the case where each voxel in an imaging medium is treated as a perturbed region and the derivative is taken with respect to $\Delta\mu_a$ and $\Delta\mu_s$, we obtain the pMC absorption and scattering Jacobians:

$$J_{\mu_a} = \sum_{i=1}^{N} l^i(r) exp\left(-\sum_r \mu_a(r) l^i(r) \right) \tag{7.22}$$

$$J_{\mu_s} = \sum_{i=1}^{N} \left(l^i(r) - \frac{k^i(r)}{\mu_s(r)} \right) exp\left(-\sum_r \mu_a(r) l^i(r) \right) \tag{7.23}$$

where N is the total number of detect photons, $l^i(r)$ is the path length and $k^i(r)$ is the number of collisions for detected photon i inside voxel r. The time dependence of the Jacobians can be determined by grouping detected photons by their arrival times and applying (7.22) and (7.23) for photon detected for each time bin.

For implementation of the perturbation approach, a temporary array is used to store the pathlength and number of scattering events in each voxel for an individual photon. Values in the array are used to contribute to the estimate of the absorption and scattering Jacobians only if the photon is detected. The temporary array is cleared before each new photon is injected into the medium. Most previous implementations of pMC for optical tomography have been for the generation of CW weight matrices (Sassaroli et al., 1998; Hayakawa et al., 2001). While (Chen &

Intes, 2009) showed a time domain implementation in a simulation study, to our knowledge experimental reconstruction of both absorption and scattering using a time-domain implementation has not been performed.

7.4.3 Comparison between Adjoint and Perturbation Monte Carlo methods

Both implementations of the time dependent Jacobian matrices for absorption and scattering are compared in this section. The fundamental difference between these two methods is that in the transport adjoint method every photon that is simulated contributes to the estimate of Jacobians while for the pMC method, only photons which hit detectors contribute to the estimate. This results in lower SNR for the pMC estimate compared to the transport adjoint estimate. To show the difference in SNR, the absorption Jacobian matrices were calculated for both methods using the same number of photons (100 million). The medium was a 2 cm thick slab with a single source and detector at the center of the slab. Contour plots of the Jacobians for both methods at the time gate corresponding to the peak of the TPSF (t = 0.4 ns) are plotted in Fig. 7.3(a). It can be seen while both Jacobian matrices have similar features, the photon noise is more apparent in the pMC estimate.

To compare the accuracy of time profiles of both methods, the perturbed measurement due to an inclusion was calculated in three ways. First, the difference of the measured signal at the surface between a homogeneous medium and a medium containing either an absorption or scattering inclusion was calculated through two Monte Carlo forward simulations. This was used as the reference measurement to compare to the two Jacobian methods. The perturbed measurement was also obtained using the pMC and transport adjoint Jacobian matrices by taking the sum of the Jacobian for voxels corresponding to the known inclusions. A comparison of the three perturbed measurements for both kinds of inclusions are shown in Fig. 7.3(b). While both methods produced temporal profiles which are similar for the absorption inclusion, the pMC method showed

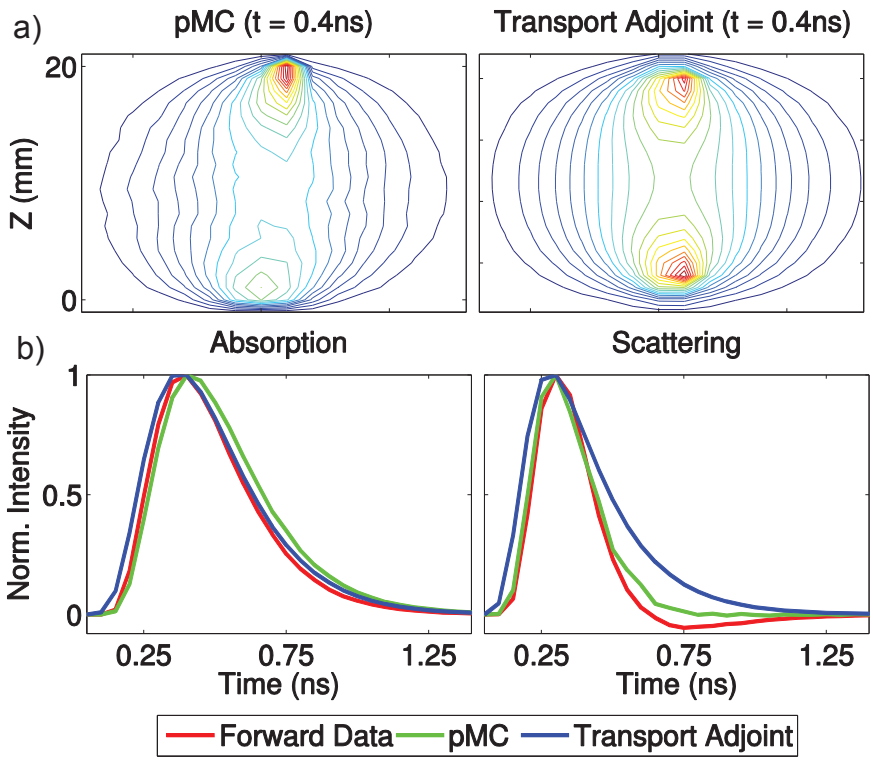

Figure 7.3: Comparison of time dependent Jacobian matrices for absorption and scattering generated using the transport adjoint and perturbation Monte Carlo (pMC) methods. (a) Contour plots of the absorption Jacobian for a single time point demonstrating the lower SNR of the pMC estimate compared to the transport adjoint. (b) Forward data due to an absorption inclusion (left) and scattering inclusion (right). Red corresponds to simulated forward data based on the photon history file, green is the pMC predicted forward data and blue is the transport adjoint predicted forward data.

a better match than the transport adjoint method to the shape of the temporal profile of the perturbed measurement in the case of a scattering perturbation. The time dependent Jacobians for simulations and experimental reconstructions in later sections were generated using the pMC method. Due to lower SNR of the pMC estimate for the Jacobian, the number of simulated photons was set to a high number (3 billion) to compensate for this effect.

7.5 Constrained Optimization Simulations

Simulations were performed to test the constrained optimization algorithm compared to a standard reconstruction algorithm. A diffuse 2 cm $\times$ 2 cm $\times$ 2 cm slab was used for both optical property and fluorescence tomography reconstruction problems. 49 sources and 49 detectors covered the z = 0 cm and z = 2 cm planes respectively. Separate background optical properties were chosen for both types of problems.

7.5.1 Absorption and Scattering

We consider reconstruction of mediums containing both absorption and scattering inclusions. A background medium with $\mu_a = 0.1\ \text{cm}^{-1}$ and $\mu_s' = 10\ \text{cm}^{-1}$ was assumed. Absorption $\Delta\mu_a = 0.2\ \text{cm}^{-1}$ and scattering $\Delta\mu_s' = 20\ \text{cm}^{-1}$ inclusions were placed at the z = 8 mm plane separated by 1 cm. Forward data was generated by taking the path history of photons and applying the Beer-Lambert law. To obtain the difference data Δy, path histories were simulated for both the homogeneous medium and the medium containing the two perturbations. Since the data contains Monte Carlo noise, additional noise was not added. Also it should be noted that the measurement data was not generated by multiplication of the Jacobian with the optical property perturbation but was instead generated using the difference of two forward simulations so that any inaccuracies due to linearization of the nonlinear forward function is included in the simulations results.

For reconstruction, the time dependent Jacobian matrices for absorption and scattering were generated using the pMC approach. Figure 7.4 shows reconstruction results for both a constrained optimization approach and a DTD approach. In the DTD reconstruction, there is significant absorption cross-talk at the location of the scattering inclusion. This cross-talk is of similar magnitude to the reconstructed value of absorption at the location of the absorption inclusion. In the constrained optimization reconstruction, the absorption cross-talk is much reduced and there

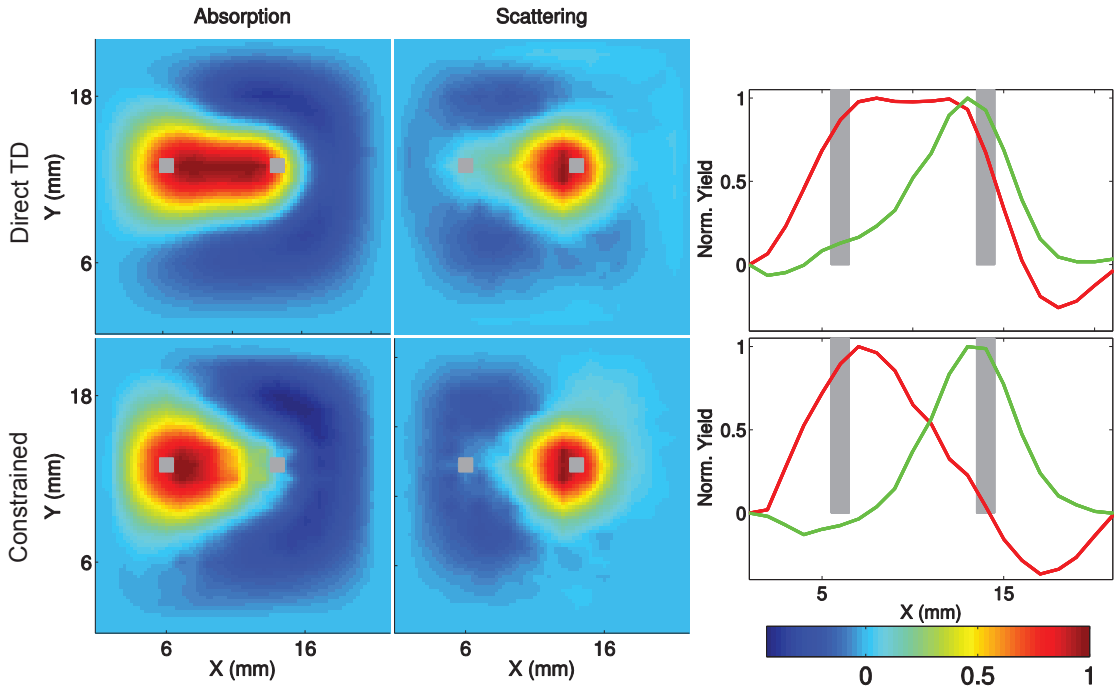

Figure 7.4: Optical property reconstruction (with a pure absorption inclusion on the left and a pure scattering inclusion on the right) using a standard DTD approach (top row) and constrained optimization approach (bottom row). XY plots are shown on the left at the plane of the inclusion (z = 8 mm). Line plots are shown on the right for both reconstruction methods with absorption in red and scattering in green.

is a clear separation between the two inclusions.

7.5.2 Lifetime Tomography of Short Lifetimes

We also consider the problem of separating fluorescent yields in the case where either the condition for asymptotic measurement of lifetimes ($\tau > \tau_D$) does not hold or the condition that the completed evolution of the background diffusive term ($t \gg \tau_{BD}$) does not hold. In both cases, an ATD approach for reconstruction cannot be applied. Using the constrained optimization approach developed in this chapter, we look to obtain solutions similar to the ATD solution.

The background medium was chosen to be $\mu_a = 0.01\ \text{cm}^{-1}$, $\mu_s' = 35\ \text{cm}^{-1}$ resulting in an intrinsic diffuse time scale, $\tau_D = 0.6$ ns. We look to separate inclusions that are 6 mm, 4 mm and 2 mm apart with lifetimes $\tau_1 = 0.60$ ns and $\tau_2 = 0.80$ ns. The lifetimes were chosen such that τ_1 does not meet the asymptotic condition and the $\tau_{effectve} = 2.4$ ns for τ_2 is sufficiently large that measurement data will not be within the asymptotic regime. The inclusions are embedded in the center of the medium at the z = 1 cm plane. For reconstruction, 16 time gates, 100 ps apart were used from the decay portion of the fluorescence TPSF.

Figure 7.5 shows the reconstruction for both methods with $\tau_1 = 0.60$ ns component in red and $\tau_2 = 0.80$ ns. component in green. It can be seen that there is significant cross-talk in the DTD reconstructions. Also, the locations of the recovered inclusions are incorrect for all separations. In the constrained optimization reconstructions, the inclusions are significantly better separated. Although the recovered location of the lifetime component for τ_1 for the 6 mm separation is shifted relative to the true position, the localization is still much better than the direct TD approach. The constrained optimization approach was also able to correctly localize the inclusions for 4 mm and 2 mm separations.

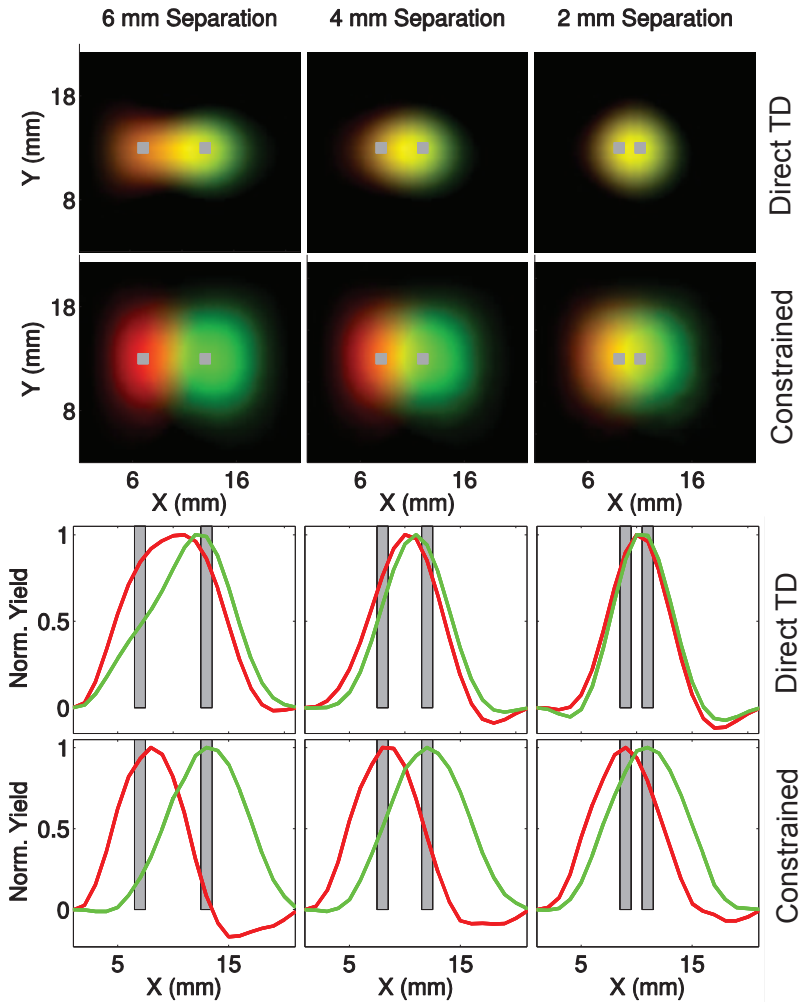

Figure 7.5: Comparison of reconstructions obtained from applying constrained optimization and direct TD approach to non asymptotic time gates. Red and green corresponds to the yield distributions for 0.60 ns and 0.80 ns, respectively. The true locations of the inclusions are shown in gray. In the first two rows, the X–Z plots are generated by assigning the recovered yields to the red (0.60 ns) and green (0.80 ns) components of the RGB colormap. The bottom two row shows line plots for constrained optimization and direct TD along the x-axis at the depth of the inclusion.

7.6 Experimental Validation

7.6.1 Phantom Creation

To verify the validity of the proposed algorithm, an optical phantom containing both absorption and scattering inclusions was constructed. Both gel and liquid based phantoms were explored for this study. Gel phantoms were created using agar powder dissolved in water with Intralipid as the scattering agent and nigrosin as the absorber. Liquid phantoms were created from water, Intralipid and nigrosin. Through preliminary measurements it was determined the gel phantoms performed better than the liquid phantoms due to two main problems associated with using liquid phantoms. First, the use of both plastic and glass tubes to hold liquid inclusions resulted in artifacts in the reconstruction that did not appear when using the gel phantom. Second, homogeneous nature of the background medium could not be maintained in the liquid phantom as the Intralipid formed aggregates in a short amount of time (< 10 minutes). The structure of the aggregates was also observed to changed with time. In many instances, it was found that the heterogeneities caused by the aggregates were of a greater magnitude than the inclusions themselves. A gel-based phantom resulted in a stable background medium for the duration of the measurement. Next, the procedure to create the gel phantom is outlined.

A gel phantom with an absorption and a scattering inclusion were made inside an 88 mm diameter dish. Agar powder (1%) was dissolved in water heated to 90°C. After the solution was cooled down to 65°C, Intralipid and nigrosin was added to achieve $\mu_a = 0.1\ \text{cm}^{-1}$ and $\mu_s' = 10\ \text{cm}^{-1}$. After mixing, the solution was poured into a dish to form the base of the phantom. Two additional solutions were made from the stock gel solution. In one, the nigrosin concentration was increased by a factor of 5 while in the other the Intralipid concentration was increased by a factor of 5. Both solutions were poured into separate smaller dishes up to a height 8 mm and allowed to solidify. Inclusions were cut from these dishes and placed on the base of the phantom

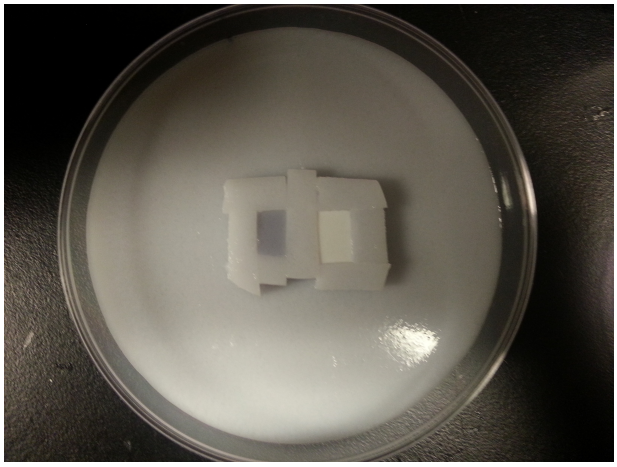

Figure 7.6: A gel phantom with absorption (left dark rectangle) and scattering (right light rectangle) inclusions placed on top of a solidified base. To complete the construction of the phantom, a stock gel solution is poured on top of the dish, covering the two inclusions.

in the large dish such that their center to center separation was 1 cm. The stock gel solution was allowed to cool to 30°C before it was poured into the large dish to cover the absorption and scattering inclusions. Prior to imaging, the phantom was removed from the dish so that the dish material would not influence the measurement. Fig. 7.6 shows a picture of the 88 mm dish with the absorption and scattering inclusions placed on top of the solidified gel, before the inclusions were covered with the background medium.

7.6.2 Measurement Methods

A tomography scan was performed using the time-domain optical tomography system described in chapter 3. Due to the sensitivity to timing jitter of optical property reconstructions, a reference laser source was incorporated into the system so that an instrument response function (IRF) measurement and the TPSF through the medium could be taken simultaneously. At each source, the IRF was compared against a reference IRF and if the norm of the difference between the two IRF's was above a certain predefined threshold, the measurement was reacquired. This procedure

reduced the amount of overall jitter in the measurements at the cost of increased acquisition time. 34 Source were split in two rows with coverage over the area of the inclusions. Adjacent sources were separated by 1.84 mm and the two rows were separated 3.68 mm. For reconstruction, detectors were chosen to be at the same x-y locations as the sources at the top surface of the phantom. The TPSF was recorded for 26 times gates with a 100 ps step size. Measurements of two samples were performed. The first measurement was the phantom containing an absorption and a scattering inclusion. The second was a homogeneous phantom created using the same stock gel solution in a separate dish.

7.6.3 Results

Fig. 7.7 shows the reconstructed absorption and scattering maps in the X-Z plane normalized to their maximum intensities. The white box in the X-Z plots marks the true location of the absorption and scattering inclusions. For the direct TD approach, it can be seen from the X-Z plots there is significant cross-talk for both absorption and scattering reconstructions. The line plots show that the amount of reconstructed absorption in the place of the scattering perturbation is 0.84 of the maximum absorption and the amount of reconstructed scattering in the place of the absorption perturbation is 0.75 of the maximum scattering. Using the constrained optimization approach, the X-Z plots show a significant reduction of the cross-talk term. The absorption cross-talk is reduced to 0.60 while the scattering cross-talk is reduced to 0.14. It can be seen that the amount of the reduction in cross-talk differs for absorption and scattering. This can be attributed to the better isolation of the absorption signal from the scattering signal in the temporal profile of their responses. Although the constrained algorithm showed a significant reduction in cross-talk for the experimental phantom, the cross-talk in simulations was found to be further reduced. This mismatch may result from different sources of modeling error including the assumption of pure absorption and scattering perturbations.

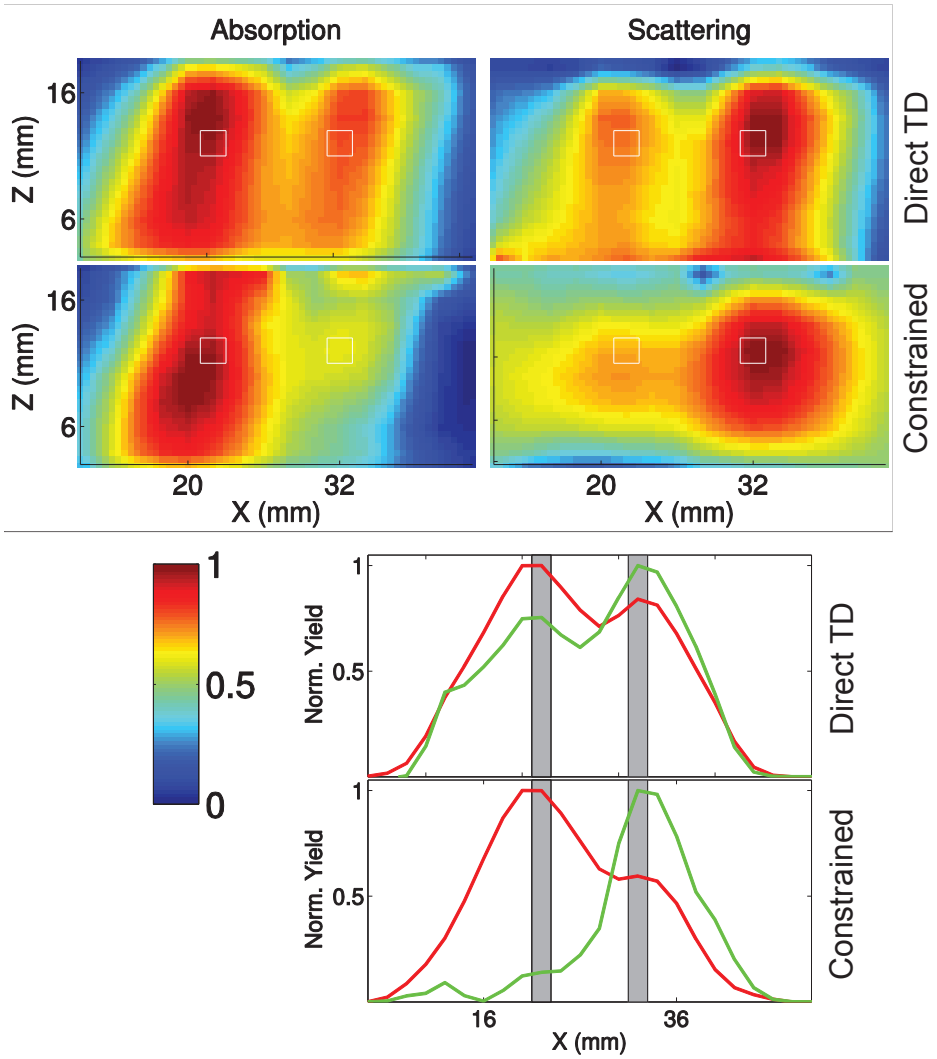

Figure 7.7: Experimental reconstruction of an absorption (left) and a scattering (right) inclusion embedded in a gel phantom. The first two rows show X-Z plots for the constrained optimization and direct TD reconstructions. The bottom two rows show line plots for both methods with absorption in red and scattering in green.

7.7 Discussion

In this chapter, we have shown a novel method for separation of multiple parameters in time domain optical tomography. The methods developed in this chapter differ from those applied in the previous chapters due to significant differences in the underlying model. For lifetime multiplexing in the asymptotic regime, the weight matrix factorization allowed intuitive methods to improve separation by allowing a fitting step to time domain basis functions unaffected by regularization. In a general time-domain multiparameter problem such as separation of absorption and scattering, this type of fit cannot be performed. This is readily apparent from in Fig. 7.8 where the normalized time profiles of the absorption/scattering Jacobian calculated using the pMC approach are shown for various voxels. While the ATD basis functions consist of the same exponential basis function for each voxel, it can be seen that the temporal basis function for absorption or scattering are voxel dependent. Despite the lack of unique basis functions, the general behavior of the absorption and scattering time profiles are still distinct with the scattering time profiles lagging behind the absorption time profiles. The methods developed in this chapter were able to exploit this difference and provide better separation of two parameters than standards methods.

Even though the constraints explored in this chapter were based on minimization of cross-talk, the presented methods should have more general applications for other types of constraints. Different performance measures are often sought for in optical tomography reconstructions. Two examples of such measures include localization of an object or spatial uniformity of the point spread function. Although these measures are related to the error in the reconstruction, like cross-talk they may not be optimized when only error is considered. Historically, achieving performance improvements for such measures involved incorporation of different types of regularization matrices. For example, the use of a spatial regularization matrix in spatially variant regularization has been shown to achieve constant image resolution throughout the medium (Pogue et al., 1999). How-

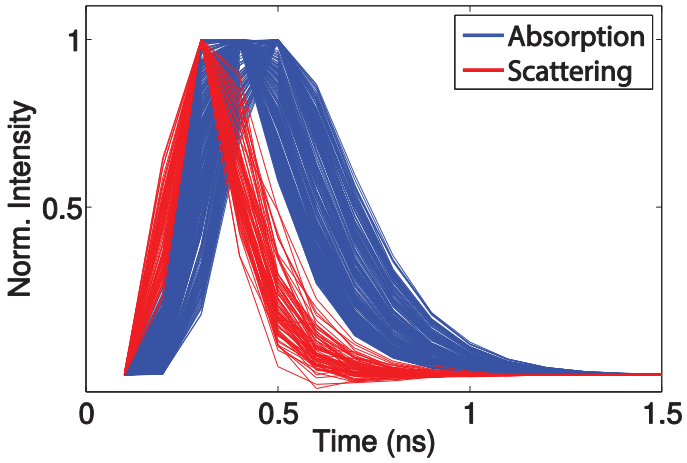

Figure 7.8: Normalized temporal profile of TD Jacobian matrices for absorption (blue) and scattering (red) at various voxels in the medium.

ever, assuming a particular measure can be framed as a suitable constraint, applying a constrained optimization approach may be advantageous since it can produce the optimal linear estimator for a given constraint. Also, unlike methods which involve incorporation of regularization matrices, the constrained optimization approach does not assume a particular form for the inverse operator. Despite any potential gains in performance of a particular measure, the increase in MSE for a given constraint should also be considered.

Any constraint on the MSE functional will increase the overall error of the reconstructed solution. In the case of minimizing cross-talk where the goal was to shape the resolution matrix, the overall sensitivity to noise will also increase. This presents a unique challenge for these types of algorithms as the constraints cannot be too strict. Otherwise, the total error will make the estimator impractical to use. More analysis is required comparing the effect of constraints on the total error to gain a better understanding of the trade-offs inherent in these methods.

8

In Vivo Lifetime Multiplexing of Organ Targeted Dyes

8.1 Introduction

Fluorescence tomography is becoming an increasingly important in vivo imaging modality for tracking the 3D distribution of specific biomarkers and providing functional information in both

small animal (Ntziachristos et al., 2005) and human subjects (Corlu et al., 2007). The ability of the technology to image multiple fluorophores simultaneously (multiplexing) is especially useful for tracking distinct but parallel biological processes in vivo. Multiplexing is commonly achieved by taking spectral measurements and unmixing the contribution of individual fluorophores by their emission spectra. However, fluorophores in the near-infrared (NIR) frequently have highly overlapping and broad spectra, requiring complex unmixing algorithms for efficient multiplexing. Fluorescence lifetime (FLT) provides an additional contrast mechanism by which fluorophores with similar spectra can be multiplexed. To image FLT, a time-resolved approach such as frequency domain (FD) or time domain (TD) must be employed. In this chapter, we will present experimental and theoretical methods for in vivo tomographic lifetime multiplexing using TD measurements. Also measurements are conducted to validate the accuracy of the technique using embedded fluorophore phantoms, fluorophores implanted in sacrificed mice and organ targeted fluorophores in living mice.

8.2 Methods

8.2.1 Fluorescence lifetime tomography

The fluorescence tomography forward problem for fluorophores with lifetimes τ_n and yield distributions $\eta_n(r)$ embedded in a diffuse medium of support Ω can be modeled as a linear equation relating the yields η_n and the time dependent measurement U at all source/detectors pairs (r_s, r_d):

$$U(r_s, r_d, t) = \sum_{n=1}^{N} \int_{\Omega} W_n(r_s, r_d, r, t)\eta_n(r)d^3r \tag{8.1}$$

where W_n is a sensitivity function expressed as the double convolution of source and detector Green's functions and the fluorescence decay term: $W_n = G^x(r_s, r, t) \bigotimes exp(-t/\tau_n) \bigotimes G^m(r, r_d, t)$.

We have shown (Kumar et al., 2006) that when $\tau_n > \tau_D$ (τ_D is the intrinsic diffuse time scale), W_n can be decomposed into a spatial and temporal term: $W_n(r_s, r_d, r, t) = \overline{W_n}(r_s, r_d, r)exp(-t/\tau_n)$ where $\overline{W_n}$ is the continuous wave (CW) reduced absorption sensitivity function. This allows the yield reconstruction to be performed using an Asymptotic TD (ATD) approach (Kumar et al., 2006) as follows. First, a set of decay amplitudes can be determined through linear fit of the asymptotic portion of the measured data using exponential basis functions. In matrix form, the amplitudes a_n^{ij} (source i, detector j) can be expressed in terms of the yield through a block diagonal matrix:

$$\begin{bmatrix} a_1^{ij} \\ \vdots \\ a_N^{ij} \end{bmatrix} = \begin{bmatrix} W_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & W_N \end{bmatrix} \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_N \end{bmatrix} \tag{8.2}$$

The yield can then be recovered from the amplitudes using standard Tikhonov regularization. We have shown that the ATD method results in significantly less cross-talk than inverting the full TD sensitivity matrix. Also, initially performing a linear fit in the time domain before applying a regularized inverse on CW weight matrices can significantly reduce the reconstruction time and memory requirements compared to a direct inversion.

8.2.2 Imaging System

A detailed description of our imaging system can be found in chapter 3. Briefly, a Ti-Sapphire femtosecond laser is used for excitation in the near-infrared. Time resolved measurements are acquired using a voltage gated image intensifier attached to a cooled CCD camera. A HRI delay unit is used to control the voltage-gating relative to the laser pulse. For 3D surface capture, a standard camera is used to take multiple white-light images centered on the mouse at various angles and a photo-stitching technique is applied using the Autodesk 123D Catch software.

8.3 Experimental Verification of Lifetime Tomography

In this section, we present three experimental measurements to show the ability of lifetime based methods to reconstruct complex shapes in a simple medium, to accurately localize the location of fluorescent inclusions in a sacrificed mouse and to reconstruct organs specific dyes in a living mouse.

8.3.1 Embedded Fluorophore Phantom

To demonstrate the ability of ATD to recover fluorescence yield distributions with complex shapes, a phantom was created by embedding a tube inside a cuvette. The cuvette was then secured to the bottom of an 88 mm culture dish and covered with a scattering medium: 1% Intralipid ($\mu_s' = 10\,\mathrm{cm}^{-1}$) and nigrosin at 4.34 µg / mL ($\mu_a = 0.1\,\mathrm{cm}^{-1}$). The cuvette acts as the background medium and contains IRDye 800cw with lifetime $\tau_1 = 0.4$ ns while the tube contains IRDye 800RS with lifetime $\tau_2 = 1.0$ ns. The phantom was imaged in transmission geometry with 104 sources below the dish and 104 detectors directly above the sources at the surface of the phantom. A schematic of the phantom is shown in Fig. 8.1.

The slice in the reconstructions corresponding to the center of the tube (z = 0.8 cm inside the scattering medium) is plotted in Fig. 8.2 for both ATD and CW. It can be seen that ATD is not only able to correctly distinguish the dye in the tube from the dye in the cuvette, the shape of the tube is also revealed in the reconstruction. On the other hand, the CW reconstruction is not able to identify the presence of an inclusion inside the background medium.

8.3.2 Embedded Dyes in Sacrificed Mouse

We next studied the ability of the TD tomography system to accurately localize inclusions with lifetime contrast in a complex medium. Two NIR dyes (IRDye 800CW, $\tau_1 = 0.5$ ns and DTTC

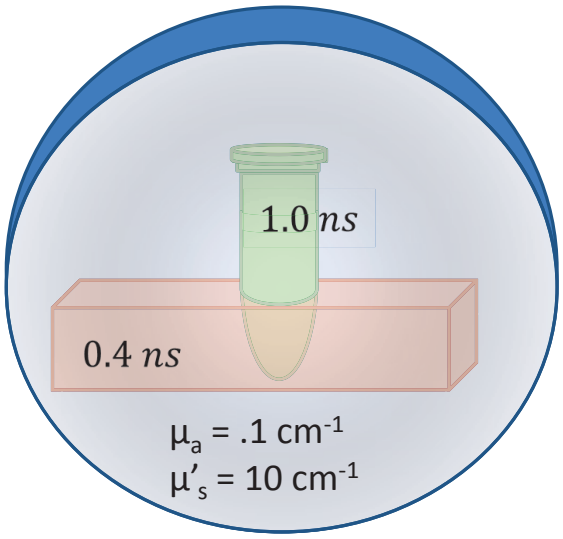

Figure 8.1: Schematic of an embedded fluorophore phantom. A tube containing a scattering medium with $\tau_2 = 1.0$ ns is embedded inside a cuvette with $\tau_1 = 0.4$ ns. Both objects are submerged inside a 88 mm culture dish.

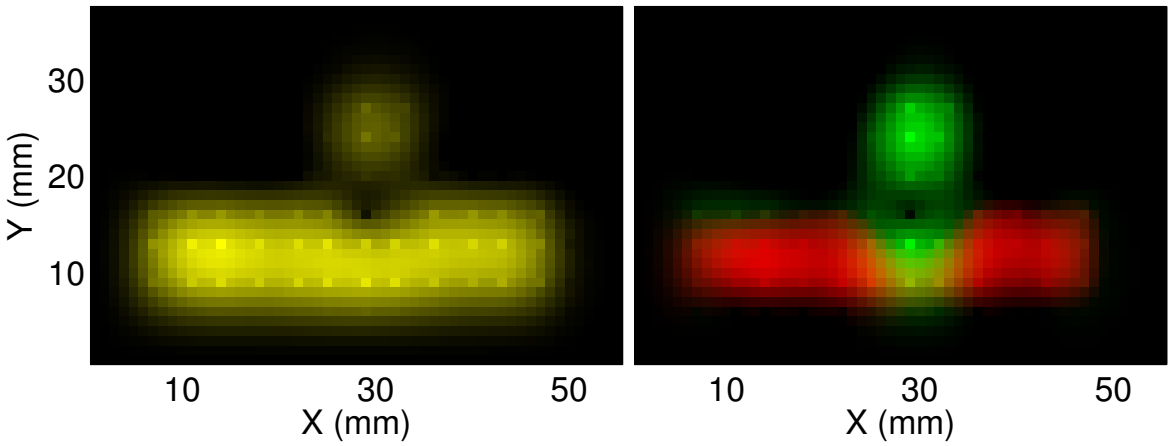

Figure 8.2: CW and ATD reconstructions for the embedded fluorophore phantom. X-Y slices are taken at the height of the center of the tube (z = 0.8 cm from the bottom of the dish). In the ATD reconstruction, the lifetime components for $\tau_1 = 0.4$ ns and $\tau_2 = 1.0$ ns are shown as green and red respectively.

in ethanol, $\tau_2 = 1.3$ ns) in the form of pellets were embedded 4 mm apart in a polypropylene tube of diameter 1 mm. The tube was implanted into the esophagus of a sacrificed mouse. Both dyes were excited at $\lambda_x = 740$ nm and emission was collected with an 800 nm long-pass filter. A planar reflection measurement of the tubes were taken before implantation. Full TD measurements were taken at 45 source positions covering the brain and neck regions of the mouse. 128 detector positions were selected in a dense grid over a wide region surrounding the sources. A CT image was taken just prior to the optical measurement.

Fig. 8.3 shows in vitro amplitude maps for the tubes showing separation of the two lifetime components. For emission data through the sacrificed mouse, the yield distributions for each dye was obtained using both an ATD approach on the TD data and an intensity based CW reconstruction. The reconstructed yields were co-registered with a CT image that shows the true locations of the tube and dyes. From Fig. 8.3, it can be seen that while the CW reconstruction results in a single distribution located in between the two dye locations, the ATD reconstruction is able to accurately localize both the short lifetime (red) and long lifetime (green) dyes. It should be noted CT was not used to provide prior information for the optical reconstruction and only used for verification of the accuracy of the localization.

8.3.3 Lifetime Multiplexing of Organ Specific Dyes in Living Mice

We next applied the system and algorithms to tomographically recover multiple organ targeting dyes in vivo. First, the lifetime distributions and organ specificity of various NIR dyes were characterized in vivo. The screened dyes included both commercially available dyes and those provided by our collaborator, Dr. Hak Soo Choi. The objective of the screening was to find dyes which have relatively narrow lifetime distributions in vivo. These dyes could then be used with the lifetime multiplexing algorithms we have developed. In the case of dyes with broad in vivo lifetime distributions, more general nonlinear reconstruction techniques would need to be employed for

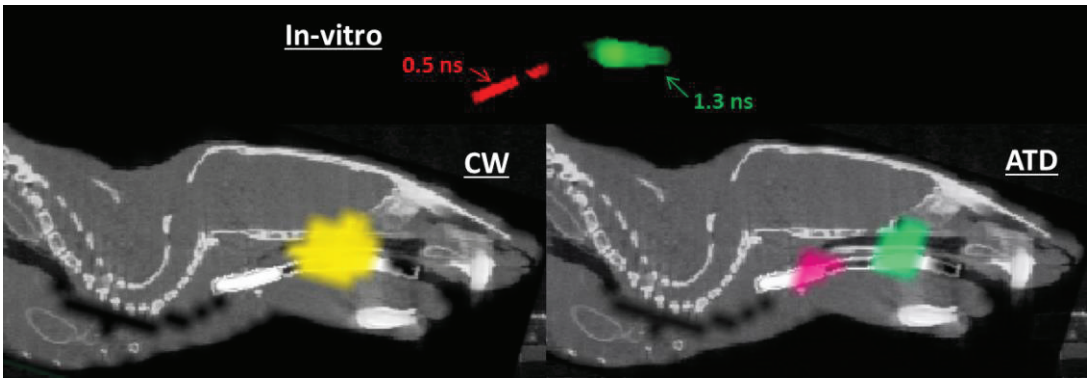

Figure 8.3: Fluorescence yield reconstruction of tube with NIR dyes (IRDye 800CW, $\tau_1 = 0.5$ ns and DTTC, $\tau_2 = 1.3$ ns) implanted in sacrificed mouse. In vitro lifetime measurement of tube is shown at top. The bottom row shows CW (left) and ATD (right) reconstructions overlaid on the CT image of mouse.

the joint recovery of lifetime and yield distributions.

Nude mice were injected with 0.1-1 mg / mL of each dye intravenously. After injection, planar fluorescence images were taken with the field of view covering the entire mouse body. Time points for multiplexing were chosen for each dye based on specificity of organ labeling, clearance time of the dye and the suggested imaging time for commercial dyes. Lifetime maps were generated at this time point for all dyes. Table 8.1 summarizes the mean in vivo lifetimes and the variances for each dye. It can be seen from the small variances in the lifetimes that the dyes displayed a nearly monoexponential behavior across the entire mouse.

For in vivo lifetime multiplexing, we consider two dyes with sufficiently separated lifetimes Osteosense 800 ($\tau = 0.53$ ns) and ZE-169-E ($\tau = 0.65$ ns). Osteosense 800 binds to bone in several hours and its fluorescence signal is detectable for days. ZE-169-E selectively targets the kidneys and achieves a strong localized fluorescence signal in 4 hours.

To demonstrate localization of multiple dyes, we simultaneously imaged both organ targeting dyes. For our measurement, a nude mouse is injected with Osteosense 800 and after 24 hours is injected with ZE-169-E. Planar images were taken 4 hours after the second injection. After

Fluorophore	Target	λ_{ex} (nm)	λ_{em} (nm)	$\tau \pm \sigma$(ps)
Indocyanine Green	-	800	835	700 ± 18
Osteosense 800	bones	780	805	650 ± 47
P800H[†]	bones	776	794	900 ± 50
ZE169-3[†]	kidneys	770	803	500 ± 22
PTN1[†]	liver	765	800	995 ± 35
Angiosense	vasculature	750	780	730 ± 30
IRDye 800cw	-	778	794	906 ± 23

[†] Dr. Choi's laboratory

Table 8.1: List of screened NIR fluorophores with their associated organ specificity, excitation/emission maxima and in vivo lifetime distributions.

tomography measurements, the mouse was sacrificed and the skin was removed. In situ planar images were then taken. In Fig. 8.4, it can be seen that amplitude maps for both in vivo and in situ measurements showed that Osteosense 800 was confined to bony structures while ZE-169-E was localized to the kidneys indicating multiexponential fits of surface measurements could delineate the structure of targeted organs.

Full tomographic reconstruction of the yield distributions for kidney and bone targeting dyes was performed in vivo. 97 sources positions, in a grid configuration were scanned from below the mouse and 319 detector positions were selected from the dorsal surface of the mouse. The ATD method using surface decay amplitudes was used for the reconstruction. Both the reconstructed yield distributions for the kidney targeting dye (green) and bone targeting dye (red) matched closely with the expected locations and known shapes of the organs being targeted (Fig. 8.5). The in vivo reconstructions also showed similarities with the Digimouse reconstructions of the same organs presented in chapter 5.

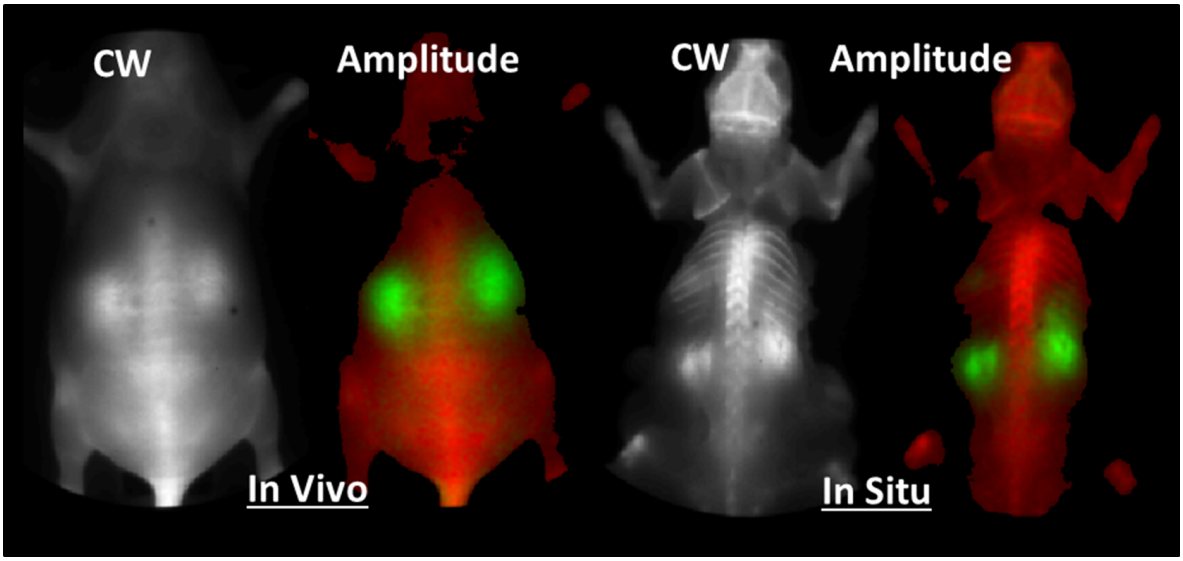

Figure 8.4: In vivo and in situ images of two organ targeting dyes. Decay amplitudes for the bone targeting dye are shown in red ($\tau = 0.65$ ns) and the kidney targeting dye in green ($\tau = 0.50$ ns). In situ images are taken with the skin removed.

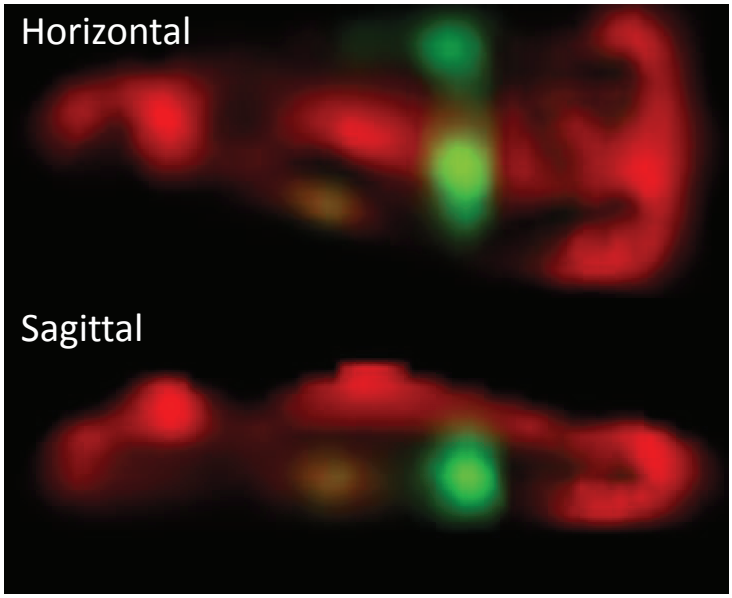

Figure 8.5: In vivo tomographic reconstruction of two organ targeting dyes. Maximum intensity projection for both the horizontal and sagittal planes are displayed. Fluorescence yield for the bone and kidney targeting dyes are shown in red and green respectively.

8.4 Summary

We have developed tomographic lifetime multiplexing methods to allow the resolution of multiple fluorophores simultaneously present in a mouse using lifetime contrast. To validate the localization accuracy for reconstructing point-like sources, we imaged tubes containing fluorescent pellets implanted in a sacrificed mouse and used co-registered CT images to verify localization accuracy. Ongoing work is focused on the imaging of organ targeted dyes with lifetime contrast in living mice. Quantitation and localization accuracy will be evaluated using ex-vivo imaging of organs as well as CT imaging. Although the present work was restricted to two fluorophores, we are exploring lifetime multiplexing of up to four fluorophores with distinct lifetimes in vivo. The methodology developed here can allow the simultaneous monitoring of parallel biological processes using multiple targeted fluorophores in mice. Future work will focus on validating localization and quantitation accuracy of the reconstruction with coregistered CT images and fluorescence lifetime images of crysosections.

9

Conclusion

In this chapter, the theoretical developments and experimental results of this dissertation are summarized. In addition, we briefly discuss some potential extensions of our work.

9.1 Chapter Summary

9.1.1 Instrumentation and Experimental Techniques

In chapter 3, a description of our imaging system was presented. Novel experimental techniques were implemented including the adaptation of photo-stitching methods for 3D surface reconstruction of mice, incorporation of a spatial filter in the form of a mechanical mask for improved SNR performance and the addition of an IRF reference measurement for jitter reduction in optical property tomography. Also, a custom Matlab based acquisition software was written and used for all experimental measurements reported in this dissertation. System calibration measurements were performed including camera flat field measurements and pixel dependent instrument response function measurements.

9.1.2 The Asymptotic TD Approach

Previously, (Kumar et al., 2005) had shown that in TD fluorescence tomography, the time scale for the diffusive and the fluorescence responses can be analyzed separately. Also, when the time constant for the decay of the fluorescence response is greater than the time constant for the decay of the diffuse response, the fluorescence term can be measured asymptotically. In the asymptotic regime, the TPSF can be reduced to a form identical to lifetime microscopy where the measurement can be written as a sum of multiexponential decay functions. Simulations using asymptotic time gates showed better separation than time gates on the rising portion of the TPSF where there is little lifetime contrast. It was also verified experimentally that asymptotic time gates could separate fluorescent inclusions 6 mm apart in a mouse shaped phantom (Kumar et al., 2008b).

We have built on this previous work in the following ways. First experimentally we showed in liquid dish phantoms that the ATD method could separate inclusions closer than the diffuse

point spread function for CW (which was found to be 5-6 mm for our set of experimental parameters). We verified ATD could even correctly localize targets with center to center separation of 1.4 mm. This was the smallest separation that was attempted and the minimum separation could be shorter. We also compared the ATD approach with the DTD approach for separation of lifetime components. We first showed the full time dependent weight matrix can be factored into a product of a basis function matrix and a spatially dependent block diagonal weight matrix. This allowed us to explicitly write the inverse operators for ATD and DTD and show they are mathematically distinct. In addition, when both operators were applied to the same asymptotic time gates, we found that ATD performed better than DTD for separation of closely spaced inclusions.

Further insights were gained into the connection between ATD and DTD when we showed that both inverse operators could be written in two forms, one form operating on the raw TD data and the other form operating on the decay amplitudes resulting from a linear fit of the raw TD data to exponential basis functions. We showed the form operating on decay amplitudes was more computationally efficient than the form operating on the raw TD data since both the DTD and the ATD solutions could be obtained from first fitting for decay amplitudes, a step that can be computed quickly for any number time gates and then inverting the CW weight matrices. We also showed that when the DTD inverse problem is interpreted using the Bayesian formalism, the ATD inverse operator can be obtained from the DTD inverse operator by replacing the actual data covariance matrix with a matrix containing terms from the exponential basis functions.

However, despite these results it was still unclear why ATD performed better than DTD for the problem of separating fluorophores. Even though, image reconstruction fidelity can be measured in a standard way using reconstruction error, to our knowledge, a general method to quantify cross-talk between different lifetime components for all voxels in a medium has not previously been investigated. We proposed to measure cross-talk for a linear inverse operator by evaluating terms on the off-diagonal blocks of its resolution matrix. This measure for cross-talk made intu-

itive sense as it represents the response of one lifetime to a true object of another lifetime. When the resolution matrices for DTD and ATD were calculated, it was found that for ATD, the basis function matrix was canceled out and resolution matrix was block diagonal while the resolution matrix for DTD was not generally block diagonal. Hence, we concluded that DTD is the estimator which produced solutions with lower reconstruction error while ATD produced solutions with higher reconstruction error but with zero cross-talk.

We have also shown results pertaining to asymptotic analysis of the decay portion of the fluorescence response. It was shown previously that the condition for asymptotic detection of lifetimes is based on a comparison between the intrinsic fluorescence lifetime and the diffuse time scale. We extend this previous analysis by introducing a new time constant based on both time scales which determines the time after the peak of the TPSF that marks the beginning of the asymptotic regime. This allows standardization of the fitting procedure for decay amplitudes.

For future work, one question which remains unanswered when DTD and ATD are interpreted using the Bayesian formalism is the error performance of the ATD estimator relative to other estimators with similar properties. Is the ATD estimator the minimum mean square error estimator with the constraint that the resolution matrix is block diagonal? We have performed preliminary simulations (not shown in this dissertation) which indicates this to be true but have not shown this result analytically.

9.1.3 Hybrid Approach for Lifetime Tomography)

In chapter 5, we consider how to utilize the entire TPSF for image reconstruction. We showed in chapter 4 that for time points in the asymptotic regime, separation of lifetimes with zero cross-talk can be achieved using the ATD method. However, there are time points prior to the asymptotic regime that cannot be used in such an approach. In addition, studies using early photon tomography have shown that reconstruction with minimally scattered photons are higher resolution than

CW reconstruction and reconstructions using late arriving photons (Niedre et al., 2008). We developed two hybrid methods to combine the resolving power of early time gates with the lifetime separation power of late time gates.

In the first hybrid method, the ATD reconstruction was used as a spatial prior for the early time gate reconstruction. In the second, the ATD and early time gate weight matrices were combined into a single weight matrix for inversion. Separate regularization was applied to the decay amplitudes and early time gates and a positivity constraint was imposed to reduce cross-talk. Simulations compared reconstructions using ATD and early time gates alone with hybrid reconstructions to show the improvements in both resolution and separation using the hybrid approaches. Simulation phantoms including point inclusions, complex shapes of varying degrees of overlap and a digital mouse atlas with different organ targets were used to validate the hybrid approaches. Finally, experimental data showed the application of the technique for separating closely spaced inclusions in a liquid phantom.

Future work would focus on demonstrating the hybrid approach in more complicated phantoms experimentally and ultimately for in vivo applications. A technique that could provide high resolution (1-2 mm) reconstructions for multiple biological processes simultaneously in a living mouse would have direct applications for many important biological problems.

9.1.4 Spectral multiplexing

Multiplexing of multiple fluorophores is more commonly performed in the spectral domain than the time domain. In chapter 6, the observation was made that both spectral and lifetime methods can be expressed as a product of a block diagonal spatial weight matrix containing CW weight matrices and a basis functions matrix. Due to this factorization, both methods allow for two types of reconstructions, one in which the full weight matrix is directly inverted and one in which the spatial weight matrix and the basis functions matrix are inverted separately with the spatial weight

matrix inverted with regularization and the basis functions matrix inverted without regularization. However, a fundamental difference between spectral and lifetime multiplexing is that the order of the spatial weight matrix and basis functions matrix product is switched relative to each other. We interpret this difference from a physical point of view. In spectral multiplexing, multiple fluorophores are combined at the voxel level and then propagated through the medium while in lifetime multiplexing, each lifetime component is first propagated through the medium and then combined at the measurement level. We also found that this difference in ordering meant that lifetime methods provided an inherent advantage in separating fluorophores over spectral methods. This was shown by calculation of the resolution matrices for both spectral and lifetime methods. For spectral methods, neither direct nor indirect reconstruction methods produced a block diagonal resolution matrix while as we showed in chapter 4, the ATD method for lifetime multiplexing does produce a block diagonal resolution matrix. This reversal in order also makes the spectral methods more sensitive to incorrect modeling of optical properties. We also showed for spectral methods, the dependence of cross-talk on the amount of variation in the optical property spectrum of the medium. To show the better separation performance of lifetime compared to spectral methods, we consider the problem of separating the relative yields of three fluorophores in a digital mouse model with optical property spectra obtained from literature. We showed that the ATD method could accurately reconstruct relative yields while the spectral method could not.

Future work would develop algorithms which combine spectral and lifetime methods into a single inverse problem. Such a method would be especially suited for multiplexing applications where some subsets of the fluorophores have unique spectral basis functions but have little lifetime contrast while other subsets have well separated lifetimes but highly overlapping spectral basis functions.

9.1.5 Constrained Optimization

It has been shown that due to non-uniqueness, absorption and scattering cannot be separated using CW data (Arridge & Lionheart, 1998). Time domain data results in better separation but a significant cross-talk component remains (Gao et al., 2002). In chapter 7, we use concepts developed in chapter 4 and apply them to general multi-parameter inverse problems. We seek to find inversion methods which produce better separation of parameters at the expense of higher reconstruction error. Previous attempts to improve separation have taken advantage of the different temporal profiles of the absorption and scattering weight matrices. Absorption is known to be more heavily weighted towards late time gates while scattering is more heavily weighted towards early time gates. In (Nouizi et al., 2011), a method using time points at the tail of the TPSF for absorption only reconstruction and time points on the rising portion of the TPSF for scattering only reconstruction was found to reduce inter-parameter cross-talk. We look to develop a more general method which can use all time gates simultaneously to reconstruct for both absorption and scattering while achieving better separation than standard techniques.

We first implemented two methods for calculation of time domain Jacobian matrices for absorption and scattering. We compared a transport equation adjoint implementation with a perturbation Monte Carlo (pMC) implementation. For single absorption and scattering inclusions, we found that the forward data predicted by the pMC approach better matched the true forward data for the inclusions than the forward data predicted by the transport adjoint approach. We next defined cross-talk in a statistical manner as the correlation between a true object of one parameter type and the reconstructed object for another parameter type. Next, we solved for the estimator that has minimum mean square error while placing constraints on the total cross-talk. We applied this estimator derived from constrained optimization to a simulation phantom which contains both an absorption and a scattering inclusion and found it shows better cross-talk performance

than a standard direct TD approach. We also apply the constrained optimization approach to the problem of lifetime multiplexing for time points not in the asymptotic regime. For inclusions 2 mm, 4 mm and 6 mm apart, we showed the constrained optimization approach could produce reconstructions with separation performance similar to an ATD approach applied to asymptotic time gates. Finally, we applied a constrained optimization approach in an experimental phantom made from agar gel, with embedded absorption and scattering inclusions and found significant reduction in cross-talk using the constrained optimization approach compared to a standard direct TD approach

While better separation of optical parameters was the goal of the present study, future work could use a similar approach for optimization of other properties of a reconstructed image. For example, it has been shown that the point spread function throughout the medium using standard inversion methods is non uniform. A constrained optimization approach could be used to find a solution which minimizes for mean squared error while placing a constraint on the spatial uniformity of the the point spread function at all voxels in the medium.

9.1.6 In Vivo Lifetime Multiplexing

In chapter 8, three multiplexing experiments were conducted to demonstrate the feasibility of lifetime multiplexing in vivo. A phantom with a complex geometry was constructed using a cuvette and tube to demonstrate the retrieval of complex shapes in a turbid medium. Full tomographic scan was performed and the asymptotic reconstructions showed recovery of the embedded cuvette while CW reconstruction could not detect existence of an embedded object. To test the localization accuracy of lifetime multiplexing in vivo, fluorescent pellets were inserted into the esophagus of a sacrificed mouse. Reconstructions showed accurate localization using ATD methods while CW reconstruction could only recover a single object between the two pellets. Finally, organ targeted dyes were screened for their in vivo lifetimes and two were chosen for multiplexing, one

targeting the bone and the other, the kidneys. Planar in vivo and ex vivo imaging showed the dyes showed good specificity to the organs of interest. Full 3D reconstruction showed recovery of lifetime components matched expected locations for the targeted organs.

9.2 Conclusion

In this thesis, we have presented various improvements in experimental technique and novel reconstruction algorithms for optical molecular imaging using fluorescence and optical property contrast. We have also shown experimental validation of the different reconstruction techniques in simple optical phantoms and in the living mouse.

References

Abdoulaev, G. S. & Hielscher, A. H. (2003). Three-dimensional optical tomography with the equation of radiative transfer. *Journal of Electronic Imaging*, 12(4), 594–601.

Alexandrakis, G., Rannou, F. R., & Chatziioannou, A. F. (2005). Tomographic bioluminescence imaging by use of a combined optical-pet (opet) system: a computer simulation feasibility study. *Physics in medicine and biology*, 50(17), 4225.

Arridge, S., Schweiger, M., Hiraoka, M., & Delpy, D. (1993). A finite element approach for modeling photon transport in tissue. *Medical physics*, 20(2), 299–309.

Arridge, S. R. (1999). Optical tomography in medical imaging. *Inverse problems*, 15(2), R41.

Arridge, S. R. & Lionheart, W. R. (1998). Nonuniqueness in diffusion-based optical tomography. *Optics Letters*, 23(11), 882–884.

Aydin, E., De Oliveira, C., & Goddard, A. (2002). A comparison between transport and diffusion calculations using a finite element-spherical harmonics radiation transport method. *Medical physics*, 29(9), 2013–2023.

Bastiaens, P. I. & Squire, A. (1999). Fluorescence lifetime imaging microscopy: spatial resolution of biochemical processes in the cell. *Trends in cell biology*, 9(2), 48–52.

Ben-Israel, A. & Greville, T. N. (2003). *Generalized inverses*, volume 13. Springer.

Berezovska, O., Ramdya, P., Skoch, J., Wolfe, M. S., Bacskai, B. J., & Hyman, B. T. (2003). Amyloid precursor protein associates with a nicastrin-dependent docking site on the presenilin 1–γ-secretase complex in cells demonstrated by fluorescence lifetime imaging. *The Journal of neuroscience*, 23(11), 4560–4566.

Bertero, M. & Boccacci, P. (1985). Introduction to inverse problems in imaging. 1998. *Bristol, UK: IOP Publishing*.

Bloch, S., Lesage, F., Gandjbakhche, A., Liang, K., Achilefu, S., & McIntosh, L. (2005). Whole-body fluorescence lifetime imaging of a tumor-targeted near-infrared molecular probe in mice. *Journal of biomedical optics*, 10(5), 054003–054003.

Boas, D., Culver, J., Stott, J., & Dunn, A. (2002). Three dimensional monte carlo code for photon migration through complex heterogeneous media including the adult human head. *Optics express*, 10(3), 159–170.

Boas, D. A. (1996). *Diffuse photon probes of structural and dynamical properties of turbid media: theory and biomedical applications*. PhD thesis, University of Pennsylvania.

Boas, D. A., Brooks, D. H., Miller, E. L., DiMarzio, C. A., Kilmer, M., Gaudette, R. J., & Zhang, Q. (2001a). Imaging the body with diffuse optical tomography. *Signal Processing Magazine, IEEE*, 18(6), 57–75.

Boas, D. A., Gaudette, T., Strangman, G., Cheng, X., Marota, J. J., & Mandeville, J. B. (2001b). The accuracy of near infrared spectroscopy and imaging during focal changes in cerebral hemodynamics. *Neuroimage*, 13(1), 76–90.

Bremer, C., Ntziachristos, V., & Weissleder, R. (2003). Optical-based molecular imaging: contrast agents and potential medical applications. *European radiology*, 13(2), 231–243.

Byrd, R. H., Lu, P., Nocedal, J., & Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16(5), 1190–1208.

Case, K. M. & Zweifel, P. F. (1967). Linear transport theory.

Chaudhari, A. J., Ahn, S., Levenson, R., Badawi, R. D., Cherry, S. R., & Leahy, R. M. (2009). Excitation spectroscopy in multispectral optical fluorescence tomography: methodology, feasibility and computer simulation studies. *Physics in medicine and biology*, 54(15), 4687.

Chen, J. & Intes, X. (2009). Time-gated perturbation monte carlo for whole body functional imaging in small animals. *Optics express*, 17(22), 19566–19579.

Chen, J., Venugopal, V., & Intes, X. (2011). Monte carlo based method for fluorescence tomographic imaging with lifetime multiplexing using time gates. *Biomedical optics express*, 2(4), 871–886.

Chen, K., Perelman, L. T., Zhang, Q., Dasari, R. R., & Feld, M. S. (2000). Optical computed tomography in a turbid medium using early arriving photons. *Journal of biomedical optics*, 5(2), 144–154.

Cheong, W.-F., Prahl, S. A., Welch, A. J., et al. (1990). A review of the optical properties of biological tissues. *IEEE journal of quantum electronics*, 26(12), 2166–2185.

Corlu, A., Choe, R., Durduran, T., Rosen, M. A., Schweiger, M., Arridge, S. R., Schnall, M. D., & Yodh, A. G. (2007). Three-dimensional in vivo fluorescence diffuse optical tomography of breast cancer in humans. *Optics express*, 15(11), 6696–6716.

Culver, J. P., Durduran, T., Furuya, D., Cheung, C., Greenberg, J. H., & Yodh, A. (2003). Diffuse optical tomography of cerebral blood flow, oxygenation, and metabolism in rat during focal ischemia. *Journal of cerebral blood flow & metabolism*, 23(8), 911–924.

Das, B., Liu, F., & Alfano, R. (1997). Time-resolved fluorescence and photon migration studies in biomedical and model random media. *Reports on Progress in Physics*, 60(2), 227.

Dogdas, B., Stout, D., Chatziioannou, A. F., & Leahy, R. M. (2007). Digimouse: a 3d whole body mouse atlas from ct and cryosection data. *Physics in medicine and biology*, 52(3), 577.

Eppstein, M. J., Fedele, F., Laible, J., Zhang, C., Godavarty, A., & Sevick-Muraca, E. M. (2003). A comparison of exact and approximate adjoint sensitivities in fluorescence tomography. *Medical Imaging, IEEE Transactions on*, 22(10), 1215–1223.

Fang, Q. & Boas, D. A. (2009). Monte carlo simulation of photon migration in 3d turbid media accelerated by graphics processing units. *Optics express*, 17(22), 20178–20190.

Gao, F., Zhao, H., & Yamada, Y. (2002). Improvement of image quality in diffuse optical tomography by use of full time-resolved data. *Applied optics*, 41(4), 778–791.

Garofalakis, A., Zacharakis, G., Meyer, H., Economou, E. N., Mamalaki, C., Papamatheakis, J., Kioussis, D., Ntziachristos, V., & Ripoll, J. (2007). Three-dimensional in vivo imaging of green fluorescent protein-expressing t cells in mice with noncontact fluorescence molecular tomography. *Molecular imaging*, 6(2), 96–107.

Godavarty, A., Sevick-Muraca, E. M., & Eppstein, M. J. (2005). Three-dimensional fluorescence lifetime tomography. *Medical physics*, 32(4), 992–1000.

Goergen, C. J., Chen, H. H., Bogdanov, A., Sosnovik, D. E., & Kumar, A. T. (2012). In vivo fluorescence lifetime detection of an activatable probe in infarcted myocardium. *Journal of biomedical optics*, 17(5), 0560011–0560016.

Graves, E. E., Culver, J. P., Ripoll, J., Weissleder, R., & Ntziachristos, V. (2004). Singular-value analysis and optimization of experimental parameters in fluorescence molecular tomography. *JOSA A*, 21(2), 231–241.

Graves, E. E., Ripoll, J., Weissleder, R., & Ntziachristos, V. (2003). A submillimeter resolution fluorescence molecular imaging system for small animal imaging. *Medical Physics*, 30(5), 901–911.

Grimm, J., Kirsch, D. G., Windsor, S. D., Kim, C. F. B., Santiago, P. M., Ntziachristos, V., Jacks, T., & Weissleder, R. (2005). Use of gene expression profiling to direct in vivo molecular imaging of lung cancer. *Proceedings of the National Academy of Sciences of the United States of America*, 102(40), 14404–14409.

Hansen, P. C. (1992). Analysis of discrete ill-posed problems by means of the l-curve. *SIAM review*, 34(4), 561–580.

Hartwig, R. (1986). The reverse order law revisited. *Linear algebra and its applications*, 76, 241–246.

Haselgrove, J. C., Schotland, J. C., & Leigh, J. S. (1992). Long-time behavior of photon diffusion in an absorbing medium: application to time-resolved spectroscopy. *Applied optics*, 31(15), 2678–2683.

Hayakawa, C. K., Spanier, J., Bevilacqua, F., Dunn, A. K., You, J. S., Tromberg, B. J., & Venugopalan, V. (2001). Perturbation monte carlo methods to solve inverse photon migration problems in heterogeneous tissues. *Optics letters*, 26(17), 1335–1337.

Holt, R. W., Tichauer, K. M., Dehghani, H., Pogue, B. W., & Leblond, F. (2012). Multiple-gate time domain diffuse fluorescence tomography allows more sparse tissue sampling without compromising image quality. *Optics letters*, 37(13), 2559–2561.

Hou, S., Rice, W. L., Hyun, H., Choi, H. S., Bacskai, B., & Kumar, A. T. (2014a). In vivo tomographic fluorescence lifetime multiplexing of organ targeted fluorophores. In *Biomedical Optics* (pp. BM3A–49).: Optical Society of America.

Hou, S. S., Rice, W. L., Bacskai, B. J., & Kumar, A. T. (2014b). Tomographic lifetime imaging using combined early-and late-arriving photons. *Optics letters*, 39(5), 1165–1168.

Hutchinson, C. L., Lakowicz, J., & Sevick-Muraca, E. M. (1995). Fluorescence lifetime-based sensing in tissues: a computational study. *Biophysical Journal*, 68(4), 1574–1582.

Hutchinson, C. L., Troy, T. L., & Sevick-Muraca, E. M. (1996). Fluorescence-lifetime determination in tissues or other scattering media from measurement of excitation and emission kinetics. *Applied optics*, 35(13), 2325–2332.

Hyde, D., de Kleine, R., MacLaurin, S. A., Miller, E., Brooks, D. H., Krucker, T., & Ntziachristos, V. (2009). Hybrid fmt–ct imaging of amyloid-β plaques in a murine alzheimer's disease model. *Neuroimage*, 44(4), 1304–1311.

Jacques, S. L. (2013). Optical properties of biological tissues: a review. *Physics in medicine and biology*, 58(11), R37.

Jares-Erijman, E. A. & Jovin, T. M. (2003). Fret imaging. *Nature biotechnology*, 21(11), 1387–1395.

Kim, A. D., Hayakawa, C., & Venugopalan, V. (2006). Estimating optical properties in layered tissues by use of the born approximation of the radiative transport equation. *Optics letters*, 31(8), 1088–1090.

Klohs, J., Baeva, N., Steinbrink, J., Bourayou, R., Boettcher, C., Royl, G., Megow, D., Dirnagl, U., Priller, J., & Wunder, A. (2009). In vivo near-infrared fluorescence imaging of matrix metalloproteinase activity after cerebral ischemia. *Journal of Cerebral Blood Flow & Metabolism*, 29(7), 1284–1292.

Klose, A. D. (2009). Radiative transfer of luminescence light in biological tissue. In *Light Scattering Reviews 4* (pp. 293–345). Springer.

Klose, A. D. & Hielscher, A. (2003). Fluorescence tomography with simulated data based on the equation of radiative transfer. *Optics letters*, 28(12), 1019–1021.

Krainov, A., Mokeeva, A., Sergeeva, E., Agrba, P., & Kirillin, M. Y. (2013). Optical properties of mouse biotissues and their optical phantoms. *Optics and Spectroscopy*, 115(2), 193–200.

Kumar, A. T., Raymond, S. B., Bacskai, B. J., Boas, D. A., Georgescu, R., Khismatullin, D., Holt, R. G., Castagner, J. L., A'amar, O., Bigio, I. J., et al. (2008a). Comparison of frequency-domain and time-domain fluorescence lifetime tomography. *Optics letters*, 33(5), 470–472.

Kumar, A. T., Raymond, S. B., Boverman, G., Boas, D. A., & Bacskai, B. J. (2006). Time resolved fluorescence tomography of turbid media based on lifetime contrast. *Optics express*, 14(25), 12255–12270.

Kumar, A. T., Raymond, S. B., Dunn, A. K., Bacskai, B. J., & Boas, D. A. (2008b). A time domain fluorescence tomography system for small animal imaging. *Medical Imaging, IEEE Transactions on*, 27(8), 1152–1163.

Kumar, A. T., Skoch, J., Bacskai, B. J., Boas, D. A., & Dunn, A. K. (2005). Fluorescence-lifetime-based tomography for turbid media. *Optics letters*, 30(24), 3347–3349.

Lakowicz, J. R. (2006). *Principles of fluorescence spectroscopy*. Springer.

Leblond, F., Davis, S. C., Valdés, P. A., & Pogue, B. W. (2010). Pre-clinical whole-body fluorescence imaging: Review of instruments, methods and applications. *Journal of photochemistry and photobiology B: Biology*, 98(1), 77–94.

Leblond, F., Dehghani, H., Kepshire, D., & Pogue, B. W. (2009). Early-photon fluorescence tomography: spatial resolution improvements and noise stability considerations. *JOSA A*, 26(6), 1444–1457.

Leblond, F., Ovanesyan, Z., Davis, S., Valdés, P., Kim, A., Hartov, A., Wilson, B., Pogue, B., Paulsen, K., & Roberts, D. (2011). Analytic expression of fluorescence ratio detection correlates with depth in multi-spectral sub-surface imaging. *Physics in medicine and biology*, 56(21), 6823.

Levenson, R. M., Lynch, D. T., Kobayashi, H., Backer, J. M., & Backer, M. V. (2008). Multiplexing with multispectral imaging: from mice to microscopy. *ILAR journal*, 49(1), 78–88.

Li, A., Boverman, G., Zhang, Y., Brooks, D., Miller, E. L., Kilmer, M. E., Zhang, Q., Hillman, E., & Boas, D. A. (2005). Optimal linear inverse solution with multiple priors in diffuse optical tomography. *Applied optics*, 44(10), 1948–1956.

Li, A., Miller, E. L., Kilmer, M. E., Brukilacchio, T. J., Chaves, T., Stott, J., Zhang, Q., Wu, T., Chorlton, M., Moore, R. H., et al. (2003). Tomographic optical breast imaging guided by three-dimensional mammography. *Applied optics*, 42(25), 5181–5190.

Li, A., Zhang, Q., Culver, J. P., Miller, E. L., & Boas, D. A. (2004). Reconstructing chromosphere concentration images directly by continuous-wave diffuse optical tomography. *Optics letters*, 29(3), 256–258.

Mansfield, J. R., Hoyt, C. C., Miller, P. J., & Levenson, R. M. (2005). Distinguished photons: increased contrast with multispectral in vivo fluorescence imaging. *BioTechniques*, 39(6), S33–S37.

Massoud, T. F. & Gambhir, S. S. (2003). Molecular imaging in living subjects: seeing fundamental biological processes in a new light. *Genes & development*, 17(5), 545–580.

Milstein, A. B., Oh, S., Webb, K. J., Bouman, C. A., Zhang, Q., Boas, D. A., & Millane, R. (2003). Fluorescence optical diffusion tomography. *Applied Optics*, 42(16), 3081–3094.

Montet, X., Figueiredo, J.-L., Alencar, H., Ntziachristos, V., Mahmood, U., & Weissleder, R. (2007). Tomographic fluorescence imaging of tumor vascular volume in mice 1. *Radiology*, 242(3), 751–758.

Mourant, J. R., Freyer, J. P., Hielscher, A. H., Eick, A. A., Shen, D., & Johnson, T. M. (1998). Mechanisms of light scattering from biological cells relevant to noninvasive optical-tissue diagnostics. *Applied Optics*, 37(16), 3586–3593.

Niedre, M. J., de Kleine, R. H., Aikawa, E., Kirsch, D. G., Weissleder, R., & Ntziachristos, V. (2008). Early photon tomography allows fluorescence detection of lung carcinomas and disease progression in mice in vivo. *Proceedings of the National Academy of Sciences*, 105(49), 19126–19131.

Niemz, M. H. (2007). *Laser-tissue interactions: fundamentals and applications*. Springer.

Nothdurft, R. E., Patwardhan, S. V., Akers, W., Ye, Y., Achilefu, S., & Culver, J. P. (2009). In vivo fluorescence lifetime tomography. *Journal of biomedical optics*, 14(2), 024004–024004.

Nouizi, F., Torregrossa, M., Chabrier, R., & Poulet, P. (2011). Improvement of absorption and scattering discrimination by selection of sensitive points on temporal profile in diffuse optical tomography. *Optics express*, 19(13), 12843–12854.

Ntziachristos, V., Ripoll, J., Wang, L. V., & Weissleder, R. (2005). Looking and listening to light: the evolution of whole-body photonic imaging. *Nature biotechnology*, 23(3), 313–320.

Ntziachristos, V. & Weissleder, R. (2001). Experimental three-dimensional fluorescence reconstruction of diffuse media by use of a normalized born approximation. *Optics letters*, 26(12), 893–895.

O'Leary, M., Boas, D., Chance, B., & Yodh, A. (1995). Experimental images of heterogeneous turbid media by frequency-domain diffusing-photon tomography. *Optics letters*, 20(5), 426–428.

O'Leary, M., Boas, D., Li, X., Chance, B., & Yodh, A. (1996). Fluorescence lifetime imaging in turbid media. *Optics Letters*, 21(2), 158–160.

O'Leary, M. A. (1996). *Imaging with diffuse photon density waves*. PhD thesis, University of Pennsylvania.

Painchaud, Y., Mailloux, A., Morin, M., Verreault, S., & Beaudry, P. (1999). Time-domain optical imaging: discrimination between scattering and absorption. *Applied optics*, 38(16), 3686–3693.

Patterson, M. S., Chance, B., & Wilson, B. C. (1989). Time resolved reflectance and transmittance for the non-invasive measurement of tissue optical properties. *Applied optics*, 28(12), 2331–2336.

Patterson, M. S. & Pogue, B. W. (1994). Mathematical model for time-resolved and frequency-domain fluorescence spectroscopy in biological tissues. *Applied Optics*, 33(10), 1963–1974.

Patwardhan, S. V. & Culver, J. P. (2008). Quantitative diffuse optical tomography for small animals using an ultrafast gated image intensifier. *Journal of biomedical optics*, 13(1), 011009–011009.

Pogue, B. W., Davis, S. C., Song, X., Brooksby, B. A., Dehghani, H., & Paulsen, K. D. (2006). Image analysis methods for diffuse optical tomography. *Journal of biomedical optics*, 11(3), 033001–033001.

Pogue, B. W., McBride, T. O., Prewitt, J., Österberg, U. L., & Paulsen, K. D. (1999). Spatially variant regularization improves diffuse optical tomography. *Applied optics*, 38(13), 2950–2961.

Prahl, S. A. (2001). Optical properties spectra - oregon medical laser clinic. http://omlc.ogi.edu/spectra/index.html.

Prahl, S. A., Keijzer, M., Jacques, S. L., & Welch, A. (1989). A monte carlo model of light propagation in tissue. *Dosimetry of laser radiation in medicine and biology*, 5, 102–111.

Raymond, S. B., Boas, D. A., Bacskai, B. J., & Kumar, A. T. (2010). Lifetime-based tomographic multiplexing. *Journal of biomedical optics*, 15(4), 046011–046011.

Rice, W. L., Hou, S., & Kumar, A. T. (2013). Resolution below the point spread function for diffuse optical imaging using fluorescence lifetime multiplexing. *Optics letters*, 38(12), 2038–2040.

Rice, W. L. & Kumar, A. T. (2014). Preclinical whole body time domain fluorescence lifetime multiplexing of fluorescent proteins. *Journal of biomedical optics*, 19(4), 046005–046005.

Sassaroli, A., Blumetti, C., Martelli, F., Alianelli, L., Contini, D., Ismaelli, A., & Zaccanti, G. (1998). Monte carlo procedure for investigating light propagation and imaging of highly scattering media. *Applied optics*, 37(31), 7392–7400.

Schulz, R. B., Peter, J., Semmler, W., D'Andrea, C., Valentini, G., & Cubeddu, R. (2006). Comparison of noncontact and fiber-based fluorescence-mediated tomography. *Optics letters*, 31(6), 769–771.

Schulz, R. B., Ripoll, J., & Ntziachristos, V. (2003). Noncontact optical tomography of turbid media. *Optics letters*, 28(18), 1701–1703.

Schweiger, M., Arridge, S. R., & Nissilä, I. (2005). Gauss–newton method for image reconstruction in diffuse optical tomography. *Physics in medicine and biology*, 50(10), 2365.

Selb, J., Joseph, D. K., & Boas, D. A. (2006). Time-gated optical system for depth-resolved functional brain imaging. *Journal of biomedical optics*, 11(4), 044008–044008.

Selvin, P. R. (2000). The renaissance of fluorescence resonance energy transfer. *Nature structural biology*, 7(9), 730–734.

Shcherbakova, D. M. & Verkhusha, V. V. (2013). Near-infrared fluorescent proteins for multicolor in vivo imaging. *nature methods*, 10(8), 751–754.

Stott, J. J. & Boas, D. A. (2002). A practical comparison between time-domain and frequency-domain diffusive optical imaging systems. In *Biomedical Topical Meetings*, volume 71 (pp. 626–628).

Swartling, J., Svensson, J., Bengtsson, D., Terike, K., & Andersson-Engels, S. (2005). Fluorescence spectra provide information on the depth of fluorescent lesions in tissue. *Applied optics*, 44(10), 1934–1941.

Turner, G. M., Zacharakis, G., Soubret, A., Ripoll, J., & Ntziachristos, V. (2005). Complete-angle projection diffuse optical tomography by use of early photons. *Optics letters*, 30(4), 409–411.

Valim, N., Brock, J., & Niedre, M. (2010). Experimental measurement of time-dependant photon scatter for diffuse optical tomography. *Journal of biomedical optics*, 15(6), 065006–0650068.

Vishwanath, K. & Mycek, M.-A. (2004). Do fluorescence decays remitted from tissues accurately reflect intrinsic fluorophore lifetimes? *Optics letters*, 29(13), 1512–1514.

Vogel, C. R. (2002). *Computational methods for inverse problems*, volume 23. Siam.

Wang, L., Jacques, S. L., & Zheng, L. (1995). Mcml—monte carlo modeling of light transport in multi-layered tissues. *Computer methods and programs in biomedicine*, 47(2), 131–146.

Wang, L. V. & Wu, H.-i. (2007). *Biomedical optics: principles and imaging*. John Wiley & Sons.

Weissleder, R. & Ntziachristos, V. (2003). Shedding light onto live molecular targets. *Nature medicine*, 9(1), 123–128.

Zacharakis, G., Ripoll, J., Weissleder, R., & Ntziachristos, V. (2005). Fluorescent protein tomography scanner for small animal imaging. *Medical Imaging, IEEE Transactions on*, 24(7), 878–885.

Zhan, Y., Eggebrecht, A. T., Culver, J. P., & Dehghani, H. (2012). Singular value decomposition based regularization prior to spectral mixing improves crosstalk in dynamic imaging using spectral diffuse optical tomography. *Biomedical optics express*, 3(9), 2036–2049.

Zimmermann, T., Rietdorf, J., & Pepperkok, R. (2003). Spectral imaging and its applications in live cell microscopy. *FEBS letters*, 546(1), 87–92.